Structure and Sequence of Formal and Postformal Thought: General Patterns and Individual Differences

Andreas Demetriou and Anastasia Efklides
Aristotelian University of Thessaloniki, Greece

DEMETRIOU, ANDREAS, and EFKLIDES, ANASTASIA. *Structure and Sequence of Formal and Postformal Thought: General Patterns and Individual Differences.* CHILD DEVELOPMENT, 1985, **56**, 1062–1091. According to Demetriou and Efklides's model, formal thought develops at 2 levels, those of strategy and tactics. This model postulates that tactics are structured in distinct capacity spheres—the relational, the experimental, and the correlational-probabilistic—which develop in a number of steps. The present study attempted to test most of these assumptions. For this purpose, 400 subjects, sampled from different age, sex, and SES groups, were examined by a battery of appropriate tasks. These tasks tested strategic, relational, experimental, and also postformal abilities. The application of factor analysis validated the existence of the relational and the experimental sphere. The strategic level was not clearly identified. However, both a late formal capacity and a postformal capacity for conceiving of possibilities and reflecting upon them were revealed. Joint application of discrimination levels analysis, and analysis of covariance, showed the various capacities developed in the steps described by the model. Sex and SES differences were also found. These results showed that the Demetriou and Efklides model could be integrated with the recent theories claiming that thought is structured in different inquiring systems, and that it develops postformally as systematic, metasystematic, and epistemic cognition. Flavell's notions of modification, inclusion, and mediation sequences were employed in order to interpret intra- and intersphere developmental relations. Individual differences were also discussed.

Research on formal thought has flourished during the last few years (see Keating, 1980; Kuhn, 1979; Meacham & Santilli, 1981; Neimark, 1979; Shayer & Adey, 1981). However, as Neimark has pointed out, "very few investigators interpret their results in terms of the four-group of INRC transformations or the 16 binary combinations of propositions" (1979, p. 64). This is understandable, as empirical evidence and theoretical analysis (see Demetriou & Efklides, 1981; also see Broughton, 1981) have shown that Inhelder and Piaget's (1958) *structure d'ensemble* is of questionable validity. Nevertheless, it is a fact that there are very few theoretical alternatives that have attempted to encompass the published empirical evidence on formal thought.

One which has, however, is the model proposed by Demetriou and Efklides (1981). According to this model, formal thought is structured in distinguishable capacity spheres that develop along a number of levels and steps or sublevels. The aim of this study, "given the need for construction of a more comprehensive theory of intellectual development" (Kuhn, 1979, p. viii), was to collect empirical evidence directly related to some of Demetriou and Efklides's theoretical postulations.

The first postulate of this model (MO), hereafter referred to as STRATAMO, is that the development of formal thought takes place on two levels: the level of strategy (STRA) and the level of tactics (TA). The strategic level is very broadly defined as "an overall problem solving orientation" (p. 23). A person, having attained this orientation, can connect logical statements interpropositionally, and can grasp their consequences with respect to their material referents. Thus, in principle, the person is capable of envisaging alternative points of view with respect to a problem. However, if the person is functioning only at this level, he or she is "naive" in being able to develop planful action directly related to the solution of complex problems,

The substantial contribution of Dr. M. Shayer, Centre for Science and Mathematics Education, Chelsea College, London, to the statistical analyses applied on the data is gratefully acknowledged. Thanks are also due to Dr. D. Rosenthal, University of Melbourne, Australia, for her helpful comments on an earlier version of this paper. Correspondence should be directed to Dr. Andreas Demetriou, Psychological Laboratory, School of Philosophy, Aristotelian University, Thessaloniki, Greece.

[*Child Development*, 1985, **56**, 1062–1091. © 1985 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920/85/5604-0026$01.00]

such as those represented by classical formal thought tasks. According to STRATAMO, the attainment of this level is a necessary but not sufficient condition for the emergence of the tactics level.

STRATAMO identifies the level of tactics with the attainment of the various Piagetian formal-operational schemata. However, it rejects the reduction of the various abilities expressed by these schemata into a single *structure d'ensemble.* Instead, the abilities are conceived of as being structured into three different spheres of thought: the experimental, the relational, and the correlational-probabilistic. This conception is based on two assumptions. First, that various particular thought abilities, directly related to specific modes of action and/or aspects of the environment, merge into more general thought capacities; and second, that these capacities can coexist in partial segregation.

The postulation of the three spheres was based on evidence indicating that performance in various formal tasks tends to be more consistent when the tasks are addressed to the same, rather than to different, spheres of abilities (see Demetriou & Efklides, 1981). According to STRATAMO, the performance patterns observed can only be accounted for by the supposition of crucial differences in the psychological constitution of the three spheres. Specifically, it was assumed that the overall relational sphere capacity is synthetic or convergent in character. This implies that the application of the various specific relational abilities on two or more problem factors, actions, or even mental operations "will result in their *reduction* to a relation which resolves them in a single mental product." On the other hand, the overall experimental sphere capacity appears to be mainly analytic or divergent. "That is, given a certain relation, [its] application will result in the *production* of the set of all possible relations in which the factors constituting the initial relation might appear" (p. 28). As to the domains of their application, the abilities of the relational sphere were considered as underlying the understanding of, and operating with, mathematical and/or physical ratio and proportionality concepts. The capacity of the experimental sphere was composed of combinatorial reasoning, hypothetico-deductive reasoning, and experimental behavior. The correlational-probabilistic sphere will not be discussed here, as it is outside the concern of this study.

STRATAMO postulates that the acquisition in the abilities of each sphere proceeds through a sequence of steps, given that the strategic capacity has been acquired. In the relational sphere, the developmental sequence postulated was as follows: *qualitative* understanding of covariation relations involved in various causal systems (e.g., a balance) → quantitative grasp of *direct* covariation (ratio) → quantitative grasp of *inverse* covariation (proportionality) → simultaneous quantitative handling of more than four factors or two relations at a time, and even integrating ratio and proportionality into a single system. In the experimental sphere, it is congruent with STRATAMO to postulate that the emergence order of the various experimental abilities is the following: simple production of combinations → more or less systematic isolation of variables → drawing correct inferences (forming hypotheses) out of the results of experimentation → exhaustive and precise testing of these inferences, in order to formulate a final valid interpretation of the phenomenon concerned.

These are the main STRATAMO postulates with regard to the structural and developmental relations of formal abilities. Thus, the primary aim of this study was to test the following hypotheses:

1. There are two discernible levels of thought functioning—the strategic, and the level of tactics. (Their characteristics are described above.)

2. At the level of tactics, the various abilities are organized into two distinct psychological entities that correspond to the overall capacities ascribed to the relational and the experimental sphere.

3. Each sphere's component abilities develop along a number of identifiable steps (as described above). Intrasphere developmental relations should be more robust than their corresponding intersphere relations. This hypothesis is based on the assumption that the sequential steps within a sphere represent the projection of one common underlying cognitive rule (or operation) onto higher and more refined levels of functioning. However, underlying rules differ among spheres. Therefore, since the developmental steps of two different spheres represent psychologically differing operations, they should exhibit loose developmental relations.

On the other hand, formal thought is not only a matter of cognitive functioning; it is also broadly acknowledged that various cultural (Dasen & Heron, 1980; Goodnow & Bethon, 1966), educational (Commons, Miller, & Kuhn, 1982; Demetriou & Efklides,

1979), social, and personal factors (Neimark, 1975; Overton & Meehan, 1982; Piaget, 1972; Shayer & Adey, 1981; Shayer & Williams, 1983) exert significant influences on the development of formal thought abilities. If this is so, it is of interest to identify the effects of such factors on the structuring and developmental patterning of formal thought abilities. For this reason, a complementary aim of this study was to test a number of hypotheses regarding the effects of sex, socioeconomic status (SES; low vs. high), and place of residence (rural vs. urban) on formal thought. On the basis of the existing literature, the following hypotheses were stated:

4. The structure and the sequencing of abilities should be the same across groups. This hypothesis can be justified on the premise that the structuring and sequencing ascribed by STRATAMO to the various abilities was based on psychological, functional, and processing principles. An X ability is amalgamated with a Y ability to form a structure only if it is similar to it with respect to mental processes (e.g., grasping of relations), and with respect to application domain (e.g., covariation of physical entities), etc. Moreover, an X ability (e.g., ratio) should necessarily precede a Y ability (e.g., proportionality) if it consists of simpler and/or mentally less loaded operations. Therefore, structure and sequence are considered by STRATAMO to be *strong universals*, in Dasen's (1977) terminology.

5. Males should perform better than females in the relational sphere. This could be ascribed to the spatial component of this sphere, and to females' reported disadvantage in this component. In the experimental sphere, no or limited differences should be expected, since this sphere is mainly composed of verbal and logical reasoning abilities (Ennis, 1978), where sex differences are much less clear-cut (see Burstein, Bank, & Jarvik, 1980; Shayer & Williams, 1983).

6. High-SES subjects should perform better than low-SES subjects from urban areas, who, in turn, should perform better than low-SES subjects from rural areas. This patterning of social class differences was documented in previous studies on formal thought (Goodnow & Bethon, 1966; Peluffo, 1964). Moreover, Hollos and Cowan (1973) obtained evidence indicating that life in an isolated rural environment negatively affects children's performance on at least some cognitive tasks. Unfortunately, the previous literature does not provide the grounds for stating more specific hypotheses about the possible differential susceptibility of the various formal thought components to SES influences.

Method

Subjects

A total of 400 subjects were tested. There were four age groups: (1) 150 ninth graders (mean age: 14-6; age range: 13-8 to 14-8), (2) 150 twelfth graders (mean age: 17-6; age range: 16-8 to 17-10), (3) 50 undergraduate university students (mean age: 20-11; age range: 19-1 to 21-7), and (4) 50 middle-aged subjects (mean age: 44-10; age range: 40-6 to 52-2).

Each of the first two groups consisted of three subgroups of subjects, formed according to the following socioeconomic criteria: (*a*) 50 subjects from agricultural families living in rural areas (i.e., small Macedonian villages with populations of between 500 and 3,000), (*b*) 50 subjects from working-class families living in the second largest city in Greece, and (*c*) 50 subjects from upper-middle-class families living in the same city. Both parents of each of the subjects of subgroups *a* and *b* had no more than primary education. At least one of the parents of each of the subjects of subgroup *c* was a university graduate. Children of professionals (university professors, doctors, lawyers) and of businessmen were overrepresented in this subgroup. However, all of these 300 subjects were attending state secondary schools. Thus, at least in principle, their educational experiences were the same.

Twenty-one of the undergraduates were science students and the rest (29) were arts students. Thus, the composition of this group is equivalent to the first two groups with respect to university preferences: secondary school education offers the subject pool for the whole range of university studies. In addition, this group of undergraduates was almost equally represented by socioeconomic class and place of residence.

The middle-aged subjects were all primary school teachers attending an in-service, 1-year training course. It should be noted here that these subjects had been teaching primary science, mathematics, and arts concepts for about 20 years, and that in this training course they were also attending classes in experimental methodology, science, arts, language, etc. Therefore, they could not be considered unfamiliar with the concepts encompassed in the tasks described below. Of course, it would have been preferable to have tested as many young and middle-aged adults as were included in the other two groups of subjects. Due to practical difficulties, how-

ever, this ideal could not, unfortunately, be achieved.

Each of the six adolescent subgroups, as well as the two other groups, was composed of 25 males and 25 females.

Tasks

Eight main paper-and-pencil formal thought tasks were given to each of the subjects tested. Four of them (chemical combinations, flexible rods, distances, and buildings; see Table 1) were first used by Tomlinson-Keasey and Keasey (1974). Of the remaining four, two (equilibrium in the balance and conservation of motion) were paper-and-pencil versions of Inhelder and Piaget's (1958) respective tasks developed by the authors. The last two (life on an imaginary planet and metacognitive analysis) were especially devised by the authors for this study. Each of these tasks comprised one or more items. These items aimed to test the various abilities constituting the overall capacity presumably tapped by the corresponding main task, and/or different expressions of these abilities. The correspondence of spheres, tasks, and items is shown in Table 1.

Experimental sphere.—Table 1 shows that the Chemicals 1 item was addressed to the ability to produce the 16 binary combinations, as the subjects were only asked to conceive of and write down all of the possible ways in which the described elements could be combined. However, the Chemicals 2 and 4 items were addressed to the ability to interpret data and draw conclusions from them, since the subjects had to combine the given information (the combinations themselves) in order to go beyond them and to grasp the causal relation between a given element (liquid 2 or 4) and a given result (the production of the red color). For example, that liquid 2 seems not to affect the production of the color because the color is obtained irrespective of liquid 2's presence or absence (1 + 3 + g) (1 + 2 + 3 + g). The only difference between the two items was that Chemicals 2 tested the ability to draw conclusions with respect to the operation of complete affirmation, whereas Chemicals 4 tested this same ability with respect to the operation of incompatibility, according to Inhelder and Piaget's (1958, pp. 293–303) analysis of the combinatorial system. Chemicals 3 and 5 were addressed to the ability to test conclusions, and corresponded one-to-one to Chemicals 2 and 4, respectively, since the subjects were asked to propose crucial tests that, if conducted, would enable them to make a final decision about the validity of the conclusions drawn as a response to Chemicals 2 and 4, for example, that the combination 1 + 2 + g could prove the neutral effect of liquid 2.

Finally, the Rods task was addressed to the isolation-of-variables ability, since the subjects were asked to propose how they could test the effect of a given variable by holding other confounding variable(s) constant. There was a difference, however, between Rods 1 and 2, on the one hand, and Rods 3 and 4 on the other. Specifically, the first set of items asked the subjects to propose tests on the effect of a defined variable (length) by holding width constant, which was also defined in the diagrams. The second set of items required the subjects to propose tests on the effect of an undefined variable (material) by holding length and width constant. Therefore, the second set of items required the subjects to generate a hypothesis by themselves ("If some of the rods were made of steel and some of brass, then . . . ") in order to guide their isolation-of-variables responses. Thus, Rods 1 and 2 required the subjects to apply the isolation-of-variables ability in order to solve a ready-made "problem," whereas Rods 3 and 4 required them to contribute to the formation of the "problem" to be solved, before applying this same ability, by proposing a hypothesis to be tested. Rods 2 and 4 were in fact very similar to Rods 1 and 3, respectively. The only difference between these pairs of items was the guidance provided to the subjects in the case of Rods 2 and 4. It was assumed that the addition of these two guided items would enable nonspontaneous users of the isolation-of-variables ability to circumvent their possible difficulties in applying this ability (see Stone & Day, 1978).

It is believed that the analysis attempted above shows that the various chemicals and rods items tapped the whole complex of abilities composing the capacity ascribed to the experimental sphere—that is, the ability to combine elements (Chemicals 1) and then to synthesize the combinations themselves in order to be able to draw conclusions from them (Chemicals 2 and 4) and to test these conclusions (Chemicals 3 and 4) by isolating variables with (Rods 3 and 4) or without (Rods 1 and 2) having the need to first put forward a hypothesis. The ability tapped by each of these items, together with their complexity as analyzed above, suggests that Chemicals 1, Rods 2 and 4, Chemicals 2 and 4, and, finally, Chemicals 3 and 5 and Rods 1 and 3 would be tapping the first, second, third, and fourth steps of the experimental sphere's development, respectively (see introduction).

TABLE 1

CORRESPONDENCE BETWEEN LEVELS AND/OR SPHERES OF FORMAL THOUGHT, TASKS, AND ITEMS

Levels/Spheres	Tasks	Items
Experimental sphere	Combination of chemicals: The four liquids (1, 2, 3, and 4) and the activation solution g were drawn and described as able to yield a red color when appropriately combined.	Chemicals 1: "Write down all of the tests that you would conduct to find the red color."
		Chemicals 2: "Note that somebody found that the combinations 1 + 3 + g and 1 + 2 + 3 + g do yield the color, whereas the combinations 2 + 3 + g and 1 + 2 + 3 + 4 + g do not yield the color. What conclusions can you draw about the influence of chemical 2 on the reaction?"
		Chemicals 3: "What further tests would you conduct to see if the conclusions were accurate?"
		Chemicals 4: The same as in 2, except that now reference was made to chemical 4.
		Chemicals 5: Question 3.
	Flexible rods: Twelve round rods forming various combinations of length and width were drawn. The rods were described as being the experimental material to be used to find out which rods are most flexible.	Rods 1: "Which comparisons can you make to find out if long rods bend more than short rods?"
		Rods 2: S was asked to indicate whether the comparison of a short and thick rod with a long and thin rod would be an appropriate test as a response to Rods 1.
		Rods 3: "What rods would you compare to find out if brass rods bend more than steel rods?" (None of the rods presented was described as being made of brass or steel.)
		Rods 4: S was asked to indicate whether the comparison of a long and thick rod with a long and thin rod would be an appropriate test as a response to Rods 3.
Relational sphere	Distances: Two road signs were shown. The first was complete: "Larissa: 94 miles; 152 kilometers." The second was incomplete: "Orestiada: . . . miles; 380 kilometers."	Distances: "Please compute the number to be put in the blank shown on the sign above, from the data you have been given."

	Buildings: A three-dimensional view of a 5 (length) × 4 (width) × 3 (height) building was presented, together with the cubic unit with which it had been built. A 3 × 2 strip of land was also shown. It was explained that a new building was to be built on this strip of land using the same cubic unit as was used for the first building, and it had to contain the same number of cubic units as the first building contained.	Buildings: "If the new building is to contain the same number of cubic units, how many stories high will the new building be when completed?"
	Equilibrium in the balance: A scale balance was presented. It had the same number (6) of holes on each side of its arm. A 3-kg and a 1-kg weight were suspended from hole 3 of the left and right side of the arm, respectively. A number of weights, ranging from 1 to 6 kg, were also drawn nearby.	Balance 1: "Would you add another weight on the left side of the arm in order to bring the arm to the horizontal position?" Balance 2: "Would you displace the weight that is on hole 3 of the left side of the arm onto hole 6 of the same side in order to bring the arm to the horizontal position?" Balance 3: "If the 6-kg weight were on hole 1 of the right side of the arm, and no weight were on the left side, which weight would you put on the left side, and on which hole, in order to bring the arm of the balance to the horizontal position?" Balance 4: "Given the weights of balance 3, write down all the possible actions you could take in order to bring the arm of the balance to the horizontal position. You may use any weight you like, and you may hang more than one weight on a hole if you want to."
Strategic level–metacognitive abilities	Life on an imaginary planet: An imaginary planet (X of Centaur) was described. It had an atmosphere completely similar to that of Earth up to a height of 12 inches above ground. Above this height the planet's atmosphere was defined as being "deleterious and destructive for life on Earth."	Planet 1: "Would it be possible for living organisms to exist on the X of Centaur?" Planet 2: "Would it be possible for living organisms to exist on the X of Centaur, and would they be able to survive the components of the atmosphere there that are deleterious and destructive for life on Earth?"

TABLE 1 (*Continued*)

Levels/Spheres	Tasks	Items
		Planet 3: "What types of living organisms (e.g., mammals, birds, reptiles, insects etc.) could live on the X of Centaur? If you think that life could be sustained on this planet, what would these living organisms look like (e.g., height, dimensions, etc.)?"
		Planet 4: "If there are intelligent organisms on the X of Centaur—intelligent organisms on Earth being humans—what do you think their various constructions (e.g., housing, cars, tools etc.) would look like?"
	Conservation of motion: A horizontal plane with a ditch in the middle was drawn. Several balls (described as being big and light, heavy and small, small and light, etc.) were resting at various points in the ditch. The rule governing the distance covered by each of the balls was stated.	Motion 1: "Is there any circumstance under which a ball would continue rolling forever?"
		Motion 2: "If all the factors which prevent motion were absent, would it be possible for a ball to roll forever?"
		Motion 3: "How long a distance and how much time would you need to perform an experiment which would prove that a ball could continue rolling forever?"
		Motion 4: "Is it really necessary to perform an experiment in order to prove that a ball could continue rolling forever?"
	Metacognitive analysis	Metacognition: "Try to discover and write down any common characteristics you think exist between the various tasks that you have solved. You should consider the thoughts that came into your mind when you were solving them, and the methods that you applied. Explain why you think two or more tasks have common characteristics, if you really believe they have characteristics that are truly in common." (Various comparisons were given in order to show the S how he should proceed, e.g., compare chemicals with rods, chemicals with balance, planet with motion, etc.).

NOTE.—The subjects were asked to explain their responses on all items.

Relational sphere.—The description of the items addressed to the relational sphere (see Table 1) shows that the simplest of them are the Balance 1 and the Balance 2 items, since they only required the subjects to decide whether the addition of another weight on the side of the arm that was already more loaded than the other side (Balance 1), or the displacement of this same weight even further from the fulcrum (Balance 2), would restore the equilibrium of the balance. Evidently, an intuitive understanding of the functioning of a balance would suffice to make a correct response to these items, given that the subjects were not asked to specify their responses quantitatively. However, quantification was indispensable if a satisfactory solution was to be provided to the Balance 3 and 4 items. These items required the subjects to specify the weights that, if used, would restore the equilibrium of the balance. Both of these items could be processed as simple ratio problems (Balance 3: another 6-kg weight suspended from hole 1 of the other side; Balance 4: addition of a 2-kg weight, also suspended from hole 3 of the right side of the arm) or as proportionality problems (Balance 3: 1 kg suspended from hole 6 on the other side; Balance 4: displacement of the 3-kg weight on the left side from hole 3 to hole 1). However, there was a difference between the Balance 3 and the Balance 4 items. The Balance 4 item required the subjects to find out and write down all the possible ways in which the equilibrium of the balance could be restored. Thus, this item would indicate whether the subjects were able to resolve the relations involved in a scheme not limited by specific weights and/or distances. The possession of this scheme would enable the subjects to handle more than four factors at a time. A response to the Balance 4 item asserting that the addition of a 6-kg weight suspended from hole 6 of the right arm could restore the equilibrium would be an indication of such a scheme. That is, the effect of a 3-kg weight (first factor) that is suspended from hole 3 of the left arm (second factor) can be compensated by the joint effect of a 1-kg weight (third factor) suspended from hole 3 (fourth factor) and another 6-kg weight (fifth factor) suspended from hole 6 (sixth factor) of the right side of the arm. Because of the way in which it was stated, the Balance 3 item could be solved by what could be called symmetric integration of weights and distances (6 to 1 = 1 to 6). Evidently, symmetric integration is a much simpler scheme. It only requires taking into account two weights and two distances at most, and it might even be solved through an intuitive grasp of the relations involved. However, this item could also evoke the complex scheme described in relation to Balance 4 for obvious reasons (e.g., a "4-kg weight suspended from hole 1 + 1 kg suspended from hole 2" response would indicate the employment of this scheme).

The distances and buildings tasks were addressed to the grasping and estimation of ratio and proportional relations, respectively. That is, the distances item required the subjects to understand that two different units of measurement (miles and kilometers) covary in the same direction. Thus, if a given distance is defined in terms of both units, then their relation could be estimated and applied with the aim of defining another distance in terms of the one unit, provided that this distance was already defined in terms of the other unit. The Buildings items required the subjects to apply the same set of operations, except that the dimensions involved covaried inversely rather than directly.

It seems apparent that the items analyzed above are directly related to the abilities ascribed to the relational sphere. That is, the ability to qualitatively understand covariation relations (Balance 1 and 2), to quantitatively grasp ratio (Distances) and proportions (Buildings) and to integrate them in a single mental scheme (Balance 3 and 4). Therefore, it was assumed (see introduction) that Balance 1 and 2, Distances, Buildings, and Balance 3 and 4 would be tapping the first, second, third, and fourth steps of the relational sphere's development, respectively.

Strategic level–metacognitive abilities.—The Planet 1 and 2 and the Motion 1 and 2 items were addressed to the ability to conceive of the possible. Their only difference was in the content in relation to which the possible could be conceived of. The Planet items were concerned with the possible existence of extraterrestrial life, whereas the Motion items were concerned with the possible conservation of motion. Moreover, Planet 1 and Motion 1 were unguided, whereas Planet 2 and Motion 2 were guided items. That is, the second set of items guided the subjects to conceive of the possible by providing them with the premises leading to this conception. Thus, the subjects merely had to extrapolate from the premises given the possibility implicated by them. In the case of the first set of items, the subjects also had to generate the premises by themselves. The two sets of items were used with the aim of differentiating between the net ability to conceive of and

accept possibilities if guided (second set), and the logical reasoning ability (negation of propositions) that usually generates possibilities. Evidently the two abilities are not identical, even though they are closely related.

The Planet 3 and the Planet 4 items concerned the ability to grasp the logically necessary consequences of the responses given to the Planet 1 and Planet 2 items—that, for instance, living organisms on the X of Centaur should necessarily be tiny creatures (Planet 3) if it was already considered that life could exist on this planet, but only up to a height of 12 inches above the ground (Planet 1 and 2). However, Planet 4 was evidently more complex than Planet 3, since it required the subjects to formulate their conceptions in a system integrating the whole complex of relations characterizing life on the X of Centaur. The Motion 3 item was also directed to the abilities referred to above; that is, it first tested the subjects' understanding of logical consistency (Motion 1 and 2: if balls could continue rolling forever, then [Motion 3] infinite time and space would be required in order to perform such an experiment). Second, it tested the subjects' ability to formulate their conceptions in an experimentally testable system ("If you create conditions of vacuum, cancel out any friction effects, etc., then you might be able to test whether a ball could continue rolling forever"). The Motion 4 item was addressed to the ability to differentiate the logically necessary from the empirically testable. That is, it required the subjects to understand that, in some cases, experiments are either unrealizable or even unnecessary, as their results can be foreseen with virtually unlimited certainty.

Finally, the Metacognition task was addressed to the subjects' ability to reflect on, analyze, and contrast their own thought processes, since they were asked to specify how they processed the tasks described above in order to define the similarities and differences among the thought processes evoked by them. The Metacognition task was similar to the tasks used by Commons, Richards, and Kuhn (1982), since it required the subjects to view the whole set of the remaining tasks used in this study as different systems that could be analyzed and reduced to more general/abstract rule systems.

The Planet 1, 2, and 3 items were purported to tap the abilities characterizing the strategic level. It was assumed that television, films, comics, and everyday discussions about the existence of extraterrestrial life would allow the subjects to escape from the data given, in order to conceive of and to elaborate alternative possibilities. The Motion 2 item was also directed to these abilities because of the support provided to the subjects. The Motion 1 item was considered more difficult because of the subjects' presumed unfamiliarity with the notion of inertia. The rest of the items in the Planet and the Motion task, and also the Metacognition task, were not directly related to the abilities with which STRATAMO is concerned. These items were included in the battery with the aim of collecting evidence about the nature of a postformal stage of thought development (Commons et al., 1982; Kitchener, 1983; Richards & Commons, 1984). Thus, evidence could be collected enabling the integration of STRATAMO with other relevant theories.

Little more needs to be said about the conception of possibilities tapped by the experimental and the relational tasks, on the one hand, and about most of the strategic-metacognitive tasks, on the other. Specifically, the experimental and the relational tasks could be defined as closed systems, and the possibilities evoked by them as conditional possibilities. That is, a universe of possible interconnections among the elements of the various items could be conceived of, completely exhausted, and divided into a set of "correct" and a set of "wrong" possibilities by the subjects, if they were able to apply the proper chain of reasoning to each item. For example, only three alternative possibilities could be evoked by Chemicals 2 or 4. That is, liquids 2 or 4 act positively, negatively, or neutrally with respect to the production of the color. Only one of these possibilities could be correct, and that one could be unequivocably defined by the application of a proper set of operations. Even in the case of the Balance 4 items, which could be considered the most complex of all of the experimental and relational items, the possible relations between weights and distances that could be conceived of could be completely defined through the methodical application of just a few rules on the data given.

However, the strategic-metacognitive items, with the exception of Motion 1 and 2, could be defined as open systems and the possibilities evoked by them as unlimited possibilities. That is, alternative conceptions with respect to a given set of elements could theoretically take any form not limited by these elements. For example, in the planet items, life on the X of Centaur could be conceived of as identical, more or less similar, or completely different (actually, is there any

way to specify the various different forms it could take?) from life on Earth. Moreover, the possibilities conceived of as a response to these items could not easily be proven "correct" or "wrong." For example, logical necessity may render experiments unnecessary (Motion 4), but, nevertheless, the acceptance of this conjecture is dependent on epistemological/epistemic predispositions (see Kitchener, 1983; Wood, 1983) that cannot be found in the data given. It is because of these considerations that the planet, the motion, and the metacognition tasks were regarded as tapping the strategic-metacognitive orientation of thought, whereas the experimental and the relational tasks were regarded as being directed to its more specific processing, that is, its tactical characteristics.

Scoring Criteria

The subjects' responses to each of the 24 items were classified into one of the following four categories: (*a*) uninformative, (*b*) concrete, (*c*) early formal, or (*d*) late formal. The criteria employed for the classification of responses into categories are summarized below, together with some examples. Nonsensical, irrelevant, or vastly insufficient responses were always placed into category *a*. Therefore, category *a* criteria do not appear below.

Experimental sphere.—Chemicals 1: (*b*) production of more than three combinations, but some sets of higher-order combinations (e.g., all the 3 × 3 combinations) are entirely missing; (*c*) production of *some* combinations from each of the four sets, 1 × 1 to 4 × 4; (*d*) production of all possible combinations, but one produced that was not the 4 × 4 combination. Chemicals 2 and 4: (*b*) simple description of the data given (e.g., "Chemical 2 is important because it is present in one of the combinations yielding the color"); (*c*) inferences are drawn about the role of the element in question, but they are not backed by sound logical argument (e.g., "I suppose 2 is not so important because it is present in only one of the combinations yielding the color"); (*d*) the inferences drawn are based on sound logical analysis (e.g., "Since there is a combination that yields the color without the presence of liquid 2, and since there is another combination that has the same result with the presence of liquid 2, it is evident that the color can be obtained without liquid 2; therefore it must be neutral"). Chemicals 3 and 5: (*b*) description of the elements involved, but no variation of the element being tested; (*c*) systematic variation of the element under consideration, but irrelevant elements are also introduced and/or the necessary controls are not exhausted (e.g., "If liquid 2 is neutral, then the combination 1 + 2 + 4 + g should not yield the color"); (*d*) realization of the necessary and sufficient controls in close correspondence to the hypotheses already stated (e.g., "If liquid 4 is negative and combination 1 + 3 + g yields the color, then the addition of liquid 4 to it should eliminate the color; if it doesn't, then liquid 4 is not negative"). Rods 1 and 3: (*b*) classification or ordering of the variables involved, but no variation of the variable under consideration (e.g., "Long and thin rods bend to a greater degree than short and thick rods, so these rods should bend to a greater degree than those ones"); (*c*) only some of the pairs proposed are correct, or some possibly confounding variables are not removed (e.g., two rods of equal length but different diameter are incorrectly paired "because they are very similar and one of them stands for the brass rods and the other one for the steel rods"); (*d*) successful application of the schema, "all other things being equal" (e.g., "Every pair of rods is similar in every respect because if the one was made of steel and the other of brass, you could see whether brass or steel rods bend more"). Rods 2 and 4: similar criteria as above, except that they were concerned with the subjects' acceptance or rejection of the faulty comparisons proposed to them.

Relational sphere.—Distances and Buildings: (*b*) unsuccessful application of addition and subtraction to the data given; (*c*) the successful solution of the problem through reduction to the unit under consideration; (*d*) the successful application and adequate explanation of the ratio (distances) or the proportion formula (Buildings). Balance 1 and 2: (*b*) correct response, but no explanation; (*c*) correct response with incomplete explanation (e.g., "No, this side is already heavier"); (*d*) integration of the forces operating on the two sides of the arm by the explanations provided. Balance 3 and 4: (*b*) "symmetric equilibration" based only on weights (e.g., another 6 kg suspended from hole 1 on the other side); (*c*) "symmetric integration" of distances and weights (e.g., 1 kg suspended from hole 6 on the other side); (*d*) resolution of the proportional relations involved in a scheme not limited by specific weights or distances (e.g., 2 kg from hole 3 on the other side).

Strategic level–metacognitive abilities.—Planet 1, 2, 3, and 4, and Motion 1 and 2: (*b*) geocentric conception of possibilities (e.g., "Ants or snakes could live on this planet because they are not tall, and move very close to

the ground"; "Well, I think the plane should be inclined a little bit"); (*c*) conception of "qualified" possibilities (e.g., "All known living beings could live on this planet, because if they know that oxygen is close to the ground, they could bend forward every so often and breathe"; "The clever beings could build underground cities"; "Well, I think balls stop for some reason, so I suppose they could roll forever for some other reason"); (*d*) ideal and/or alternative possibilities conceived on the basis of sound logical arguments (e.g., "Either the atmosphere of Earth defines life, or life is a phenomenon that could occur anywhere in the universe, and in any form, since different adaptation mechanisms could exploit any atmospheric conditions. In the first case, no living beings could exist on this planet; in the second case, they could exist in any form, and their constructions could be similar to, or different from, ours, according to the needs they have"; "If no factor preventing motion is in action, then balls should certainly roll forever"). Motion 3 and 4: (*b*) concrete conception of the role of experimentation and ascription of "magical" possibilities to it (e.g., Motion 3: "Few hours and one or two kilometers would suffice to show that balls could roll forever"; Motion 4: "Experiment is always necessary, because we should not believe anything, unless it has been experimentally proven"); (*c*) understanding that the form an experiment takes should be consistent with the idea being tested (Motion 3: "You need infinite time and space to prove that a ball will roll forever"), and an infinite understanding that knowledge is obtainable through means other than experimentation, since experiments cannot be performed under certain circumstances (Motion 4: "If you need infinite time and space, then you do not need the experiment, because you cannot do it"); (*d*) understanding that the experiment could be an idealized representation of natural phenomena (Motion 3: "A little bit of space and a fraction of a second would suffice, given that you have created conditions under which no factor preventing motion is in action; in this case, the speed of a rolling ball should be unchanged between any two points"), even though under some conditions it is completely unnecessary to test a premise, as its truth or falsity is based on logical assumptions (Motion 4: "You don't need an experiment because you already know for sure, that if no factor preventing motion is present, balls will roll forever").

Metacognitive analysis: (*b*) reference to the phenomenal relations between tasks (e.g., "These are science tasks, and these are math tasks"); (*c*) reference to the operations involved in the solution of the tasks (e.g., "Chemicals and rods are similar, because they require you to combine data and to test a conclusion; distances and buildings are similar, because they require you to find out a covariation relation"); (*d*) analysis of the operations involved, and interconnection of tasks pertaining to different spheres (e.g., "Chemicals and balance are similar, because they both require the combination of elements, in order to reveal their not apparent but possible relations; however, they are also different in that chemicals require combinations to find out the role elements play in them, or to test a conclusion, whereas balance requires combinations to find out how distances and weights covary; in this respect balance is similar to building, because they both require you to specify a relation in order to use it as a means of decoding the relation connecting other terms," etc.).

One hundred of the 400 protocols were evaluated independently by the authors, and their percentage agreement for each of the 24 items was calculated. This evaluation showed that the criteria presented above were very reliable. Mean interrater agreement was very high—95.8% (SD = 3.9%)—and ranged from 84.6% (on the Chemicals 3 item only) to 100% on seven out of the 24 items. Differences were resolved by discussion and each item assigned to a final category. For the purposes of statistical analysis of the data, the numbers 0, 1, 2, and 3 were assigned to the four categories, respectively.

Procedure

All subjects were tested in groups. A booklet containing the eight tasks was given to each subject. The instructions were written on the first page of the booklet and were also explained verbally by one of the authors, who was always present during the test. The instructions emphasized that (*a*) the test was not related to school or university examinations, and (*b*) the subjects could proceed through the battery at their own pace (no time limit was imposed), but they had to follow the order in which the items were presented in the booklet. This limitation was necessary, because presentation order of the tasks was randomized across subjects. The metacognition task was always presented last so that subjects could discuss the relations among the other tasks. Each of the items was written on a separate page. This precaution was taken to minimize the chance of interference between items, particularly in the case of "guided" and "unguided" items (see Table 1). For this rea-

son, the "guided" items were always presented after the other items comprising the various tasks.

Results and Discussion

The results are presented in two parts. In the first part, the results relating to the structure, the developmental sequencing, and the relations within and between sequences (i.e., those related to hypotheses 1, 2, and 3) are presented. In the second part, the results relating to the various group effects on these parameters (i.e., those related to hypotheses 4, 5, and 6) are presented.

The General Characteristics of the Structure and Development of Abilities

Structure.—To test the hypotheses concerned with the structure of abilities, factor analysis was applied to the scores representing mean performance attained by the subjects on items that presumably tapped the same ability through a different content, and/or through varying degrees of support provided to the subjects. The units thus formed are shown in the first column of Table 2. Means, rather than raw scores, were factor-analyzed, because, according to Humphreys (1976), this is the only way of significantly reducing the effects of bias variance that is associated with the different single measures of the same attribute. Moreover, this approach was adopted because, according to Thurstone (1957), the insertion of two or more measures of the same performance in the analysis may result in the creation of doublet factors that confuse, rather than elucidate, the factorial structure of abilities tapped by the tests. However, in order to ensure that the items pulled together were really different expressions of the same ability, their zero-order correlations were calculated. These correlations were very high; seven out of 10 were higher than .58 (mean $r = .63$; SD = .14).

Also, it should be noted that, first, the principal factoring method without iteration was employed. Second, the various principal component matrices obtained were submitted to an oblique rotation, since it was assumed that the factors would not be fully orthogonal. Third, the direct oblimin criterion was employed to simplify the factorial structures to be obtained. This method was adopted, rather than other more directly confirmatory ones such as the LISREL method (Joreskog & Sorbom, 1978), for three reasons. First, STRATAMO, and also the present state of our knowledge in general, does not enable one to state specific predictions about the exact contribution of each variable to the variance accounted for by a particular factor. Thus an analysis such as LISREL would seem unnecessary. Second, principal factoring without iteration is the only method enabling the exact calculation of factor scores. A major part of this study was based on analyses carried

TABLE 2

THE FACTOR STRUCTURE ABSTRACTED FROM THE PERFORMANCE OF THE WHOLE SAMPLE (N = 400)

	FACTORS			
	I	II	III	h^2
Chemicals 1	.732[c]	.292	.251	.685
Chemicals 2 + 4	.801[c]	.214	.242	.747
Chemicals 3 + 5	.805[c]	.246	.302[a]	.800
Rods 1 + 2	.725[c]	.477[a]	.014	.753
Rods 3 + 4	.677[b]	.478[a]	.068	.691
Distances	.498[a]	.538[b]	.073	.543
Buildings	.472[a]	.595[b]	.098	.587
Balance 1 + 2	.346[a]	.765[c]	.041	.708
Balance 3 + 4	.384[a]	.786[c]	.049	.769
Planet 1 + 2	.376[a]	.284	.835[c]	.919
Planet 3 + 4	.290	.283	.821[c]	.838
Motion 1 + 2	.138	.653[b]	.521[b]	.716
Motion 3 + 4	.219	.646[b]	.520[b]	.748
Metacognition	.416[a]	.422[a]	.320[a]	.453
Direct % contribution of factors to the variance of variables	21.9	18.6	12.7	57.3

NOTE.—b and c were considered as primary loadings, and a as a secondary loading. The direct % contribution of factors to the variance of variables was estimated from the factor pattern matrix according to the procedures described by Harman (1967).

out on factor scores (see below). Therefore, this method was considered preferable to others that do not enable the direct and exact estimation of factor scores. Third, the method employed here provided exact information about the degree of the variance of each particular unit analyzed that was accounted for by the factors abstracted.

Table 2 shows the three oblique factors that have eigenvalues greater than 1, which were abstracted from the performance of the whole sample of 400 subjects. These almost equipotent factors accounted for 57.3% of the total variance. The first factor correlated moderately (.40) with the second. However, the correlations between the first and the third (.18) and the second and the third factor (.23) were low.

The first factor (21.9% of the total variance) loaded primarily on the Chemical and the Rod units. Moreover, this factor accounted directly for 53.6% of the total variance of the five experimental units, but only for 4.3% of the total variance of the rest of the nine units. Evidently, this factor tapped the capacity that determined performance on the experimental and not on the other units. Therefore it may be called the "experimental capacity factor."

The second factor (18.3% of the total variance) loaded primarily on the four units (Distances, Buildings, and the two Balance units) addressed to the relational abilities. Moreover, this factor accounted directly for 35.6% of the total variance of the four relational units, but only for 9.7% of the total variance of the rest of the 10 units. Thus, it appears that this factor abstracted the overall capacity that determined performance on the relational units. Therefore, the second factor may be called the "relational capacity factor."

The third factor (12.7% of the total variance) loaded primarily on the Planet and the Motion units addressed to the abilities enabling one to conceive of ideal possibilities and to formulate them in a consistent system, to feel logical necessity and to grasp its empirical consequences, and to differentiate between the logically necessary and the empirically testable. Moreover, this factor accounted directly for 39.3% of the total variance of the four Planet and Motion units, but only for 2% of total variance of the rest of the 10 units. Therefore, this factor abstracted the capacity underlying performance on the strategic-metacognitive units, and it may be called the factor of the "capacity to conceive of possibilities and reflect upon them."

Finally, it should be noted here that some units were also primarily nested under another factor besides the one under which they should theoretically be nested (Distances, Motion 1 and 2, and Motion 3 and 4). Evidently, this finding implies the existence of abilities pertaining to more than one sphere, and this will be elaborated on later in the discussion.

Developmental sequences.—Shayer's (Shayer, 1978; Shayer & Adey, 1981; Shayer, Demetriou, & Prevez, 1983) analysis of discrimination levels was applied to the performance attained by the whole sample, on each of the 24 items, in order to trace the sequencing of the abilities represented by them. Briefly, the determination of the discrimination level of an item is a three-stage process. First, each item is assigned to one of the broad categories that are theoretically conceived of as levels of a developmental scale (e.g., intuitive → concrete → early formal → late formal). At this stage, the items' assignment to levels can either be theoretically guided, or based on the findings of other studies (e.g., according to STRATAMO, Chemicals 3 and 5, for instance, should be late formal items). Second, the performance of each subject across the whole set of items is then inspected, and the subject is placed in the highest level in which he or she succeeds in two-thirds of the corresponding items. Intermediate levels can be formed at this stage, so as to include the subjects who do not clearly fall into any one of the primary categories. That is, if a subject succeeds in only one-third of the early formal items, he or she is placed into a new level, which presumably comes before the early formal level. Finally, the percentage of subjects in each level who succeeded in each item is calculated. The discrimination level of a given item is defined as that level at which 67% of the subjects who were assessed at that level succeeded. Where a deviation is observed between the initial assignment of items to levels and the final objective discrimination level arrived at, the whole procedure is repeated until the hypothesized item levels and the subject levels become mutually consistent. In this study, the criterion of success adopted was an early formal response to each of the items. This lenient criterion was adopted, rather than a stricter one (i.e., late formal), in order to induce the earliest *formal* appearance of each of the abilities tapped by the various items scaled in relation to the earliest formal appearance of the abilities tapped by the other items. Therefore, it is clear that this method

TABLE 3

THE DISCRIMINATION LEVEL OR STEP OF THE 24 ITEMS IN RELATION TO THE SPHERE UNDER WHICH THEY NESTED

	SPHERES		
LEVELS AND STEPS	Experimental	Relational	Possibility-conceiving Reflective
Concrete level		Balance 1 and 2	Motion 2
First formal step	Chemicals 1; Rods 1, 2, and 4	Distances	...
Second formal step	Chemicals 2 and 4; Rods 3	Buildings	Planet 2
Third formal step	Chemicals 3 and 5	Balance 3 and 4	Planet 1; Motion 1 and 3
Postformal level	...	...	Planet 3 and 4; Motion 4; Metacognition

yields objective scales of the items analyzed. These scales can be considered developmental dimensions, having at their one pole the readily and commonly present abilities, and at their other pole the complex, rarely exhibited abilities. This method was used in this study because it has a number of psychometric advantages over the more conventional methods, such as scalogram analysis, which was used by other authors (e.g., Kofsky, 1966) with the aim of identifying developmental sequences (see Shayer et al., 1983).

The results of this analysis are summarized in Table 3. In this table, the various items are separated into three columns, according to their primary factor nestings. In the experimental-capacity factor, the nine corresponding items isolated three steps of formal thought functioning.[1] That is, the first manifestation of the experimental sphere was confined to an ability that enables the subjects merely to produce combinations (Chemicals 1), and to isolate variables, given either that the problem being tested was simple, or that relevant support was provided (Rods 1, 2, and 3). It is only at the second formal step that subjects proved capable of interconnecting the combinations in order to form valid working hypotheses (Chemicals 2 and 4) or of spontaneously applying the isolation-of-variables ability, after having themselves grasped the possible presence of an extraneous variable that should also be included in the controls to be effected (Rods 3). Finally, at the third formal step, the subjects proved capable of successfully planning and effecting the necessary and sufficient tests of the previously formulated hypotheses.

The six relational items discriminated between a concrete level and three formal steps of thought functioning. That is, the relational capacity manifested itself as a concrete-like ability to understand covariation relations (Balance 1 and 2) qualitatively. The ability to estimate ratio (Distances) and proportions (Buildings) quantitatively corresponded to the first and the second formal step, respectively. Finally, it was only at the third formal step that the subjects proved capable of integrating ratio and proportionality in a system that enabled them to solve the problems that required the simultaneous application of both of these abilities (Balance 3 and 4).

The eight items subsumed under the third general factor, together with the metacognition task, spread over a concrete level, two formal steps, and a postformal level of thought functioning. Thus, the general capacity represented by this factor appeared at first as a concrete-like ability to grasp an ideal possibility, such as the conservation of motion, once the premises leading to this idea have already been given to the subjects (Motion 2). However, it was only at the second formal step that the subjects managed to apply this same ability, when their familiar experience ran contrary to an ideal conception (Planet 2). At the third formal step, the subjects succeeded in conceiving of ideal possibilities, after having generated the necessary premises themselves (Planet 1 and Motion 1), and also in grasping the logical consequences of their conception (Motion 3).

[1] From now on the term "level" refers to main stages (concrete, formal, postformal), and the term "step" refers to the substages comprising a given stage.

However, it was only at a postformal level of functioning that the subjects managed to build complete complementary or competing systems (Planet 3 and 4), to differentiate between the logically necessary and the empirically testable (Motion 4), and to reflect successfully on their own thought processes (Metacognition).

Inter- and intrasphere developmental relations.—Froman and Hubert's (1980) prediction analysis was applied on four-way contingency tables (see scoring) in which the performance of the whole sample of subjects was cast. Froman and Hubert maintained that this analysis "provides a method for the direct evaluation of a specific staging proposition that takes into account the degree of confirmation that could be expected under independence" (1980, p. 137). In practice, the evaluation of a staging proposition is effected by having to define which cells of the contingency table are "error cells," that is, "those cells that should contain no observation if the proposition is perfectly reflected in the contingency table" (1980, p. 138). The proportion of the frequencies observed in the error cells to the proportion expected under the assumption of statistical independence is then calculated according to the procedures described by Froman and Hubert, and thus the value of the success of prediction is obtained.

The application of prediction analysis to the results of this study aimed to test whether the developmental sequences presented above were consistent *within* and *across* spheres. Therefore, cross-tabulations were prepared that paired performance on items belonging either to the same, or to a different, sphere. According to the third hypothesis stated in the introduction, the value of predictions concerned with sequential or synchronous acquisition of abilities pertaining to the same sphere should be considerably higher than the corresponding values concerned with abilities pertaining to different spheres.

The results obtained were in congruence with this hypothesis, as is evident from the exemplary cross-tabulations shown in Table 4. It can be seen that only 4.5% of the cases fell within the error cells (in italics) of cross-tabulation (1) that paired performance on items belonging to the same (experimental) sphere. However, 28.9% of the cases fell within the corresponding cells of cross-tabulation (2) that paired performance on items belonging to different (experimental and possibility-conceiving-reflective) spheres. In all, only a mean of 8.03% (SD = 3.04%) of the intrasphere pairings fell within the error cells, as contrasted to the mean of 29.8% (SD = 2.7%) of the intersphere pairings. Thus, the value of the various predictions tested with respect to intrasphere development was very high (mean $\hat{\nabla}$ = .602, all p's < .001). In intersphere development, no consistent priority or synchronicity relations were revealed by this analysis (mean $\hat{\nabla}$ = .113, all p's > .01), even though there was a tendency for the lower-step abilities of the one sphere to be acquired before the higher-step abilities of the other spheres.[2]

In other words, the performance on items belonging to the same sphere was patterned in an orderly way so that, in the case of presumably sequential development, the subjects solved the items addressed to earlier-coming abilities on a level higher than the level attained on items addressed to later-coming abilities (see cross-tabulation 1 of Table 4). Similarly, in the case of presumably synchronous development, the subjects solved both items at the same level. Theorists (Flavell, 1972; Wohlwill, 1973) and researchers (Jamison, 1977; Smedslund, 1964) of cognitive development considered this evidence a strong indication of the presence of direct functional relations between the abilities represented by the items analyzed—that is, that the later-coming ability stems directly from that already acquired through a process of transformation and/or generalization (sequential development), or that both abilities are expressions of the same underlying cognitive rule or operation (synchronous development). However, the performance on items belonging to different spheres was not patterned in an orderly way, so that the level attained on the one item did not condition in any systematic way the level attained on the other item (see cross-tabulation 2 of Table 4). This evidence was considered to indicate that the items represent different underlying capacities that are not functionally connected in any direct way.

Age, Sex, and SES Influences on Structure and Development

Influences on structure.—In order to identify possible age, sex, and SES influences on the structure of abilities, a separate factor analysis was applied to the performance of three age groups (across sex and SES; the per-

[2] The computer program for applying prediction analysis to data has been developed by Dr. P. Macheras, Department of Statistics, Aristotelian University of Thessaloniki.

TABLE 4

BIVARIATE FREQUENCY DISTRIBUTION BETWEEN (1) CHEMICALS 1 VERSUS CHEMICALS 5 ($\hat{\nabla}$ = .697, p < .001) AND (2) CHEMICALS 1 VERSUS MOTION 1 ($\hat{\nabla}$ = .052, p > .05)

	CHEMICALS 5			
CHEMICALS 1	a	b	c	d
a	85	*4*	2	*0*
b	53	23	*4*	*0*
c	57	35	14	*8*
d	28	41	39	7
	MOTION 1			
	a	b	c	d
a	7	63	8	*13*
b	15	44	*5*	*16*
c	6	68	14	*26*
d	4	56	15	40

NOTE.—The letters a, b, c, and d refer to the noninformative, concrete, early formal, and late formal categories, respectively. The prediction tested in both cases was that Chemicals 1 should be solved earlier than Chemicals 5 or Motion 1. The weight given to error cells (in italics) was always equal to 1.

formance of the last two age groups was combined for this analysis, because there were only 50 subjects in each), to each sex (across age and SES), and to each SES (across age and sex). The eight oblique factor structures obtained are summarized in Table 5.

Most of these factors were almost identical across groups and in close correspondence with the factors of the overall analysis. Specifically, Factor II of the 15-year-olds and Factor I of females, males, and high-SES subjects almost matched Factor I of the overall analysis that was named as the experimental capacity factor. Furthermore, Factor III of the combined last two age groups and of females, Factor IV of males and low-SES subjects, and Factor II of high-SES subjects resembled Factor II of the overall analysis that was named as the relational capacity factor. Finally, Factor III of the 15- and 18-year-old subjects, Factor I of the last two age groups, Factor II of females and males, Factor V of low-SES subjects from rural areas, and Factors II of low-SES and III of high-SES subjects were very close to Factor III of the overall analysis that was named as the factor of the capacity to conceive of possibilities and reflect upon them. Therefore, the generality of the overall factors was reasonably well supported, even when total variance was split into parts presumably subjected to differing influences.

It is true that in some cases an overall factor was divided into two group-specific factors (e.g., Factors I and IV of the 15-year-olds could be matched with the relational, whereas Factors I and IV of the 18-year-olds and II and IV of the last two age groups combined, respectively, could be matched with the overall experimental factor). This division could not be considered to reflect true structural differences, insofar as the units nested under the various factors did not produce any new combinations, compared to the nestings of the overall factors. In general, the division reflected the fact that a general capacity is composed of a number of different, more specific, abilities. Specifically, the examples noted above reflected the fact that the relational capacity comprises the ability to process mathematically expressed ratio and proportionality problems (Distances and Buildings) and the ability to grasp and process ratio and proportionality relations within a system of physical equilibrium (the Balance units); or that the experimental capacity comprises the isolation-of-variables ability (the Rod units) and the ability to formulate and test conclusions (the Chemical units).

Moreover, the division of the overall factors might well be ascribed to the relatively low number of subjects (between 100 and 200) included in the various groups. According to Guilford (1954), at least 200 subjects are required in order to obtain clear results from

TABLE 5

SUMMARY OF THE OBLIQUE FACTOR STRUCTURES ABSTRACTED FROM THE PERFORMANCE OF EACH AGE, SEX, AND SES GROUP

	A. AGE												B. SEX								C. SES											
	15-Year-Olds				18-Year-Olds				21- and 45-Year-Olds				Females				Males				Rural					Low				High		
UNITS	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	V	I	II	III	IV	I	II	III
Chemicals 1	a	c	a	a	c					b	a		c	a	a	a	b	a	b	a	a	a		b			a	c	a	c		a
Chemicals 2 + 4		c	a		c							c	c				a	a	c					c		a		c		c		
Chemicals 3 + 5		c	a		c							c	c			a	a		c					c	a	a		c	a	c		
Rods 1 + 2	b	c		a	a			b		c			c		a	b	c			a			c			c		a	a	c	a	
Rods 3 + 4	b	b		a	a	a		b		c			b		a	b	c						c			c			a	b	a	
Distances	a	b	a	a				b		a	a		a		a	b	b	a		a	b		a			b		a	a	b	a	
Buildings	b						a	b		b	b		a		a	b	b	a		a	b		a	a		b			a	a	b	
Balance 1 + 2				c		c			a	a	c				c		a			c	a	c				a		a	c	a	c	
Balance 3 + 4				c		c		a			c	a			c	a	a	a		c		c	a	a		a		a	c	a	c	
Planet 1 + 2		a	c				c		c			a	a	c				c	a					a	c		c					c
Planet 3 + 4		a	c	a	a		c		c			a		c			a	c	a		a				c		c			a		c
Motion 1 + 2	b		b	b		b	a		b		b	a		b	b	a		b		a	c					a	b		b		c	
Motion 3 + 4	c		a	a		b	b		c					a	a	b	a	c		a	c		a		a	a	a		b		c	
Metacognition	c				a		a	a	b	a	a					c	a	a				a	a	a		c					b	

NOTE.—a = .3–.49; b = .5–.69; c = .7 and above.

factor analysis. Nevertheless, the two interpretations are not incompatible. That is, the restriction of variance that is associated with the relatively low number of subjects included in the various groups did not allow factor analysis to abstract the most general dimensions underlying performance on the units analyzed. Thus, only subdimensions were abstracted by the analysis of each separate group. This interpretation could also explain the fact that the correlations between the various group-specific factors were generally lower than the correlations between the overall factors already reported. However, it should be noted that the correlations between factors representing different abilities of the same capacity, like the examples reported above, were higher (mean r = .30) than the correlations between factors representing different capacities (mean r = .22).

Nevertheless, some differences observed among the various group-specific factor structures may indicate structural or developmental differences that should be taken into account. Specifically, the decreasing number of factors observed as one proceeds from the rural to the high-SES subjects (see panel C of Table 5) requires special attention. Moreover, the height of the correlations between factors varied among groups. The mean correlations for the 15-year-olds, the 18-year-olds, and the last two age groups combined were .28, .19, and .25, respectively. The mean correlations for females and males were .22 and .30, respectively. Finally, the mean correlations for the rural, the low-SES, and the high-SES subjects were .17, .24, and .24, respectively. These points will be elaborated on later in the discussion.

Influences on development.—The analysis of discrimination levels was separately applied to the performance of each age (across sex and SES), each sex (across age and SES), and each SES group (across age and sex). The aim of these separate analyses was to specify the possible existence and the nature of intergroup differences in the developmental patterning of abilities. It should be emphatically stressed that no deviation was observed between the results of the present analyses and the results of the overall analysis that was applied on the whole sample of 400 subjects. Therefore, the relative scaling of the abilities represented by the 24 items did not seem to be affected by age, sex, or SES. Thus the results presented below are only concerned with intergroup differences in the rate of acquisition of each of the three main capacities.

For the sake of the above ends, the results of a series of analyses of covariance were used jointly with the estimation of the mean percentage of subjects who succeeded in providing early formal (or above) responses to the set of each sphere's items that discriminated at each of the developmental steps already described. It was believed that this approach would make the results of the analyses of covariance more meaningful, as it permits the translation of a significant intergroup difference in mean-scores attainment into information showing the developmental level or step attained by each group in relation to the others.

The various analyses of covariance were applied on the factor scores attained by each subject on each of the three factors originally abstracted. Factor scores, rather than raw scores, were analyzed because of methodological and theoretical considerations. Specifically, factor scores are considered as the best weightings of partially correlated tests. Thus, if a given factor is interpreted as a dimension or variable representing an underlying capacity, the factor score attained on this factor by a given subject can be considered the best composite estimate of his or her quantitative attainment of the capacity represented by the factor. Raw scores are ambiguous indexes of this capacity because they confound more than one factor (Guilford, 1954, p. 524). Moreover, the simultaneous comparison of various groups with respect to both the factor structure abstracted from their performance and the factor scores attained on the factors comprising this structure enables one to parsimoniously decide whether there are structural and/or quantitative differences among the groups compared (e.g., Buss, 1974).

Thus, three three-way age (15- vs. 18-year-olds) × sex (females vs. males) × SES (rural vs. low-SES vs. high-SES) analyses were applied on the factor scores attained by the adolescents of Factors I, II, and III (analyses Ia, IIa, and IIIa, respectively). In addition, three two-way age (18-year-olds of high-SES vs. 21-year-olds vs. 45-year-olds) × sex (females vs. males) analyses were applied on the factor scores attained by a subgroup of adolescents and the two older groups on each of the three factors (analyses Ib, IIb, and IIIb, respectively). This second set of analyses was deemed necessary by the considerably lower number of subjects included in the last two age groups. The inclusion of the 18-year-old high-SES subjects in these analyses aimed at connecting the adolescents' performance with

that of the two older age groups. In each of these six analyses, the other two factors were treated as covariates in order to control for intergroup variance on the factor tested, which could be ascribed to the other two factors. This approach seems justified by the fact that the three factors were correlated, though at a low or moderate level. The mean factor scores to which these analyses were applied are shown in Table 6. The corresponding mean raw scores attained on the units nested primarily under the three factors are also shown in this Table for indicative purposes.

Age effects: Experimental capacity.—In Factor I, which represented the experimental capacity, the effect of age was highly significant in the case of both analyses Ia, $F(1,286) = 10.019$, $p < .0001$, and Ib, $F(2,142) = 11.054$, $p < .0001$. The first of these results indicated that the performance of the 15-year-olds ($\bar{X} = -.47$) was significantly lower than the performance of the 18-year-olds ($\bar{X} = -.02$). The decomposition of the second result by means of t-contrasts indicated that the 21-year-olds ($\bar{X} = 1.024$) performed better than the high-SES 18-year-olds ($\bar{X} = .280$), $t(98) = 4.88$, $p < .001$, who attained the highest level of performance among the adolescents (see Table 6). However, the performance of the 45-year-olds ($\bar{X} = .445$) was significantly lower, $t(98) = 4.21$, $p < .001$, than that of the 21-year-olds, and not significantly different from that of high-SES 18-year-olds. These results reflected clearly the step up to which the experimental capacity was attained by the four age groups—namely, if the two-thirds success criterion is considered to indicate that a given group has attained an X capacity at an X level (Shayer, 1978), then it should be accepted that, overall, the 15-year-olds were functioning at the concrete level with respect to the experimental capacity. Only 42% of these subjects solved the first formal-step experimental items at a formal level. The 18-year-olds had nearly attained the first formal step, as 58% of them formally solved these items. The 21-year-olds were functioning at the second formal step, as 68% of them provided formal solutions to the corresponding items. Nevertheless, these subjects as a group could not be credited with the third formal-step experimental abilities, since only 48% of them provided formal solutions to the experimental items tapping the abilities of this step. The 45-year-olds were functioning at the first formal step, as the two-thirds success criterion was exceeded by 69% of them on only the respective items.

Relational capacity.—The results were quite similar with respect to the relational capacity. The age effect of IIa, $F(1,286) = 37.194$, $p < .0001$, and IIb, $F(2,142) = 3.635$, $p < .03$, analysis was significant. The first result reflected an extended improvement of performance from the age of 15 ($\bar{X} = -.45$) to the age of 18 ($\bar{X} = .23$). In terms of step attainment, their difference indicated that the 15-year-olds were functioning at the concrete level, since 74% of them understood the covariation relations involved in the two concrete items, but only 50% of them succeeded in solving at a formal level the distance(s) item that discriminated at the first formal step. The 18-year-olds were clearly functioning at the first formal step (87% of them succeeded), and they were also close to the second formal step, as 53% of them solved the respective item. In the second result, contrasts t-tests showed that the 21-year-olds ($\bar{X} = .683$) did not perform significantly better, $t(98) = 1.61$, $p > .05$, than the high-SES 18-year-olds. That is, 76% and 58% of the 21-year-olds formally solved the second and third formal-step items, respectively. However, the performance of the 45-year-olds ($\bar{X} = .004$) was significantly lower, $t(98) = 3.73$, $p < .0001$, than that of the 21-year-olds. These subjects were evidently functioning at the second formal step, as 64% of them formally solved the corresponding item, but only 45% of them solved the third formal-step items.

The capacity to conceive of possibilities and reflect upon them.—The development of the capacity to conceive of possibilities and reflect upon them appeared to be differently affected by age, even though the age effect was significant in the case of IIIa, $F(2,286) = 17.305$, $p < .0001$, and IIIb analysis, $F(2,142) = 25.486$, $p < .0001$. That is, the first result indicated that the performance of the 18-year-olds ($\bar{X} = -.09$) was *lower* than that of the 15-year-olds ($\bar{X} = .15$). This finding suggested, if nothing else, that performance based on this capacity did not improve between the fifteenth and the eighteenth year of age, as both groups were functioning at only the concrete level (77% and 67% of the 15- and 18-year-olds solved the concrete item, respectively). In the second result, contrasts t-test showed that the 21-year-olds ($\bar{X} = .598$) performed significantly better, $t(98) = 2.46$, $p < .02$, than the high-SES 18-year-olds ($\bar{X} = .139$). The 21-year-olds were clearly functioning at the second formal step, as 80% of them solved the corresponding item. The third formal step was not far beyond their capacity (55%). Nevertheless, they could not be cred-

TABLE 6

MEAN FACTOR SCORES AND STANDARD DEVIATIONS (in Parentheses) ATTAINED BY THE VARIOUS AGE, SEX, AND SES GROUPS ON THE THREE FACTORS REVEALED BY THE OVERALL FACTOR ANALYSIS

AGE, SEX, AND SES	FACTOR I		FACTOR II		FACTOR III	
	Factor Scores	Raw Scores	Factor Scores	Raw Scores	Factor Scores	Raw Scores
15-year-olds:						
Females:						
Rural	−.876 (.766)	.656 (.853)	−.699 (.750)	.920 (.952)	−.229 (.658)	.896 (.587)
Low	−.501 (.873)	.868 (.900)	−.664 (.928)	.850 (.934)	.190 (.862)	1.004 (.702)
High	−.271 (.898)	1.132 (.935)	−.275 (.907)	1.255 (1.007)	.586 (1.060)	1.388 (.851)
Males:						
Rural	−.922 (.653)	.532 (.853)	−.692 (.802)	.965 (.903)	−.308 (.975)	.868 (.761)
Low	−.207 (.915)	1.176 (1.061)	−.119 (.750)	1.475 (1.047)	.197 (1.111)	1.256 (.785)
High	−.066 (1.136)	1.284 (1.085)	.029 (1.157)	1.600 (1.374)	.477 (.846)	1.416 (.800)
18-year-olds:						
Females:						
Rural	−.632 (.627)	.708 (.771)	.183 (.553)	1.925 (.930)	−.410 (.732)	1.024 (.657)
Low	−.025 (.936)	1.300 (.934)	−.082 (.915)	1.520 (1.032)	−.173 (.830)	1.128 (.756)
High	−.007 (.802)	1.400 (.963)	.022 (.992)	1.495 (1.022)	.029 (1.025)	1.260 (.712)
Males:						
Rural	−.428 (.748)	.920 (1.017)	.216 (.886)	1.930 (.897)	−.340 (1.067)	1.072 (.855)
Low	.430 (.965)	1.648 (1.067)	.296 (.983)	1.915 (1.029)	.102 (.919)	1.384 (.982)
High	.567 (.882)	1.720 (1.013)	.729 (.836)	2.240 (.805)	.249 (.604)	1.610 (.647)
21-year-olds:						
Females	.943 (.722)	2.076 (.842)	.317 (.883)	1.935 (.984)	.303 (.782)	1.512 (.650)
Males	1.104 (.495)	2.176 (.708)	1.049 (.865)	2.475 (.733)	.892 (1.144)	1.952 (.853)
45-year-olds:						
Females	.361 (.766)	1.504 (.952)	−.539 (.803)	1.555 (1.046)	−.706 (.925)	.804 (.670)
Males	.529 (.740)	1.724 (.982)	.529 (.660)	2.275 (.822)	−.859 (.762)	.980 (.678)

NOTE.—The raw mean scores and standard deviations (in parentheses) attained by these groups on the units primarily nested under the respective factors are also shown.

ited with the abilities of the postformal level, as only 44% of them provided satisfactory responses to the postformal items. It should be emphasized here that, with respect to metacognitive analysis abilities, only 33% and 10% of the subjects, considered as functioning at the third formal step on all three spheres, gave answers satisfying the early and late formal criterion, respectively. These results corresponded closely to those reported by Commons et al. (1982) for their formal subjects who attained the postformal stages of systematic (23%) and metasystematic thinking (11%). Finally, according to contrasts t-tests, $t(98) = 5.48$, $p < .0001$, the performance of the 45-year-olds ($\bar{X} = -.783$) was much lower than that of the high-SES 18-year-olds. That is, these subjects were barely functioning at the concrete level, since only 58% of them provided satisfactory solutions. At the other levels, only 42%, 20%, and 9% of these subjects provided formal responses to the items discriminating at the second and third formal steps, and at the postformal level, respectively.

The metacognitive analysis task.—It may be of interest here to indicate the nature of the responses given on the metacognition task by the subjects operating at the postformal level. These subjects grouped the other seven tasks in categories more or less coinciding with their factor nestings—namely, almost half (44%) of them considered the Chemical and the Rod tasks to be similar, because "they both require you to combine some elements with other elements, control irrelevant variables and draw conclusions out of complex data." The distance task was said to be similar to the Building task by most (65%) of the subjects, because "they are both concerned with estimation of analogies, direct or inverse." Moreover, these two tasks were related to the Balance task (32%), because "they are concerned with the application of mathematical estimation on physical parameters, such as distances, volume and weight." Finally, 83% of the subjects identified the Planet with the Motion task, because "they are addressed to imagination and they ask you to conceive of the unknown on the basis of inadequate data." It should also be mentioned that a non-negligible number of subjects considered Rods and Balance (26%) as related, because "they require combinations, estimation of analogies and compensations, and require you to take into account the operation of physical energy." In addition, Chemicals and Balance were related by 38% of the subjects, because "both of them require you to take into account many parameters at the same time." No more than 9% of the responses placed any other combination of tasks in the same category. Therefore, it can be concluded that the *subjective structure* of formal thought that was shown by the subjects' metacognitive responses reflected its objective structure that was revealed by the factor structure of the performance itself. Such a conclusion lends credibility both to the results of factor analysis and to the megacognitive responses of the subjects.

Sex effects.—The effect of gender difference(s) on the three capacities is clear. Gender difference(s) were significant in the case of the IIa, $F(1,286) = 9.277$, $p < .003$, and IIb analyses, $F(1,142) = 30.903$, $p < .0001$, which were concerned with the relational capacity. These results indicated that, overall, the performance of males ($\bar{X} = .26$) was higher than that of females ($\bar{X} = -.26$) only with respect to the relational capacity. In other words, males were at all ages functioning at one or two steps ahead of females. Specifically, the 15-, 18-, 21-, and 45-year-old males were credited with first, second, third, and third formal-step abilities, respectively, whereas their female counterparts were credited with concrete level, first, second, and first formal-step relational abilities, respectively.

SES effects.—The impact of SES was much more extensive than that of gender difference(s). It was significant in the case of Ia, $F(2,286) = 15.914$, $p < .0001$ ($\bar{X} = -.715$, $-.076$, $.056$), and IIIa analyses, $F(2,286) = 6.462$, $p < .002$ ($\bar{X} = -.322$, $.079$, $.335$), and marginally significant in the case of IIa analysis, $F(2,286) = 2.675$, $p < .07$ ($\bar{X} = -.248$, $-.217$, $.126$, for rural, low-SES, and high-SES subjects, respectively). Contrasts t-tests showed that the performance of the high-SES subjects was higher than the corresponding performance of the rural subjects in the case of the experimental capacity, $t(198) = 6.36$, $p < .0001$, the relational capacity, $t(198) = 2.77$, $p < .006$, and the capacity to conceive of and reflect upon possibilities, $t(198) = 5.23$, $p < .0001$. However, the performance of the high-SES subjects was significantly better than that of the low-SES subjects only in the relational capacity, $t(198) = 2.39$, $p < .02$, and in the capacity to conceive of and reflect upon possibilities, $t(198) = 1.96$, $p < .05$. Finally, the performance of the low-SES subjects was better than that of the rural subjects in the experimental capacity, $t(198) = 5.29$, $p < .0001$, and in the capacity to conceive of and reflect upon possibilities, $t(198) = 3.15$, $p < .0002$. These results reflected the fact that, where there was a statistically significant difference, the sub-

jects who attained higher factor scores were functioning at one step above the rest. That is, high-SES subjects were functioning at the first and second step of the relational capacity and the capacity to conceive of and reflect upon possibilities, respectively, from the age of 15 onward. However, it was only at the age of 18 that the low-SES subjects moved to the corresponding steps, although both groups attained the first formal-step experimental abilities at this age. The rural subjects were functioning at the concrete level, even at this age, as no more than 41% of them formally solved any first formal or higher-step item. It should be stressed here that the sex and SES differences found were very consistent ones, as no one of the interactions tested ever approached significance.

General Discussion

The Structure of Abilities

The results of this study provided satisfactory support for the hypotheses about the structure of formal thought that were derived from STRATAMO. Specifically, the performance of the whole sample was factorially structured into three distinct factors. This finding is in accordance with the first two hypotheses. The three factors indicated that the ability to go "beyond the information given" (Bruner, 1974), and to envisage alternative, logically consistent points of view (Factor III), should be discerned from the experimental (Factor I) and the relational capacity (Factor II). Evidently, these results completely validated the second hypothesis, which is concerned with the differentiation of the so-called level of tactics into an experimental and a relational sphere. However, they only partially supported the first hypothesis, since Factor III could not be fully identified with the fundamental capacities ascribed by STRATAMO to the strategic level.

If that had been the case, Factor III would have been the first factor in terms of traditional theories of intelligence. Moreover, its component abilities would have been acquired *before* the acquisition of the abilities represented by the other two factors. Nevertheless, none of these requirements was met. Thus, the third factor appeared to be revealing an autonomous capacity that is rooted in a preformal stage of development (see the concrete item), but that also follows its own developmental trajectory, since it is mainly a late formal and postformal attainment.

As a result, two questions concerning structure need to be answered. First, what is the nature of the capacity revealed by the third factor, compared to the nature of the capacities revealed by the first two factors? Second, were there any findings in this study that could be considered indicative of STRATAMO's strategic level?

Churchman's (1971) concept of "inquiring systems," recently introduced by Wood (1983) to the study of cognitive development, seems particularly relevant to this first question. That is, the problems loaded on the third factor can best be considered as requiring the application of a dialectical/Hegelian, or at least a Kantian, inquiring system, in order to be solved at a satisfactory level. That is, a system that was contrary to the data bases (the dialectical/Hegelian inquiring system), or at least contrary to alternate but complementary representations (the Kantian inquiring system), had to be constructed and then synthesized or integrated. If the Planet task is taken as an example, it was necessary to realize the possibility that life in the universe cannot be defined by the forms it takes on a specific planet (e.g., Earth), or by the atmospheric conditions prevailing on Earth (i.e., a given proportion of oxygen in relation to other gases). Thus, it was necessary to pose two theoretical representations of the problem ("oxygen *as found* on Earth is the *only* precondition of life and its forms," vs. "life is *independent* of specific atmospheric conditions, since different adaptation mechanisms can make it possible—in varying forms—under completely different atmospheric conditions"). The failure to view the problem from both of these perspectives would result in a single-minded answer (i.e., life on the imaginary planet either does not exist, or, at best, the living creatures existing on that planet could not be higher than 1 foot, as oxygen is confined to this height), which would indicate that the problem was treated as one of the Leibnizian type. In fact, this was the case with the failing subjects in this study.

However, in order to view the problem from both of the above perspectives and to elaborate on them adequately, "epistemic cognition," in Kitchener's (1983) terms, seems to be needed. Epistemic cognition refers to cognitive processes that enable a subject to decide about the limits, the certainty, and the criteria of knowing, and also to "identify and choose between the form of solution required for different problem types" (Kitchener, 1983, p. 226). In turn, epistemic cognition is built upon metacognition that enables a subject to identify his cognitive processes and strategies, and to match them with the processing

demands posed by each particular task so as to arrive at an acceptable solution. For example, one might think that Balance 4 asks for a series of quantitative judgments about the possible combinations of weights and distances that would restore equilibrium. Thus, one might go on to think that two approaches in conjunction are required: first, a combinatorial approach to generate alternative weights-and-weights and/or weights-and-distances combinations; and second, the application of the related ratio and proportionality formulae in order to test that the combinations conceived of for each of the two sides would produce equal forces so as to keep the balance in equilibrium. The performance on the Metacognition task used in this study showed that even these processes were very infrequent, probably because they require postformal systematic and metasystematic operations (Commons et al., 1982). More simply, they require operations that enable the subject to specify the constituents of systems (in this study, the other seven tasks could be considered as the systems to be contrasted) in order to grasp their similarities and dissimilarities. Given these facts, the capacity to conceive of possibilities and reflect upon them could not be the first factor.

Other inquiring systems are needed to represent the relational and the experimental-capacity spheres. The Leibnizian inquiring system is the most appropriate one to model the relational sphere, since the solution of the items subsumed under the respective Factor II could be achieved by the proper application of a relevant deductive algorithm (e.g., a previously learned or developmentally constructed algebraic formula that is applicable to proportionality problems). The experimental sphere could be considered primarily a Lockean inquiring system, since the solutions to experimental problems—even though anticipated hypothetically by means of deductive, that is, Leibnizian, inquiry—can only be inductively attained on the basis of empirical evidence.

In answer to the second question posed above, one could consider that the mutual exchange of loadings between the three general factors, but mostly between the respective group-specific factors subsequently described, indirectly indicates the presence of STRATAMO's strategic level. According to these findings, each sphere appeared to have, on the one hand, some *nuclear* or *core* abilities, which are inexchangeable with the nuclear abilities of the other factors. The nuclear abilities characterizing the experimental, the relational, and the conceiving-reflecting sphere were those represented in the various Chemical, Balance, and Planet units, respectively. These units were never primarily subsumed under the same factor. On the other hand, each sphere also appeared to have some other, factorially *floating* abilities that could be primarily combined not only with their own but also with the nuclear abilities of the other spheres. The floating abilities of the experimental, the relational, and the conceiving-reflecting sphere were those represented in the Rod, Distances and Building, and Motion units, respectively. It is not unreasonable to consider these factorially floating abilities indicative of the nature of STRATAMO's strategic capacity. This means that experimental (Lockean) or conceiving-reflecting (Kantian or dialectic/Hegelian) processing of problems requires a minimum of relational (Leibnizian) processing of the elements under consideration, and vice versa. For example, the simple production of combinations may require the person to consider, at the beginning of the production attempt, how the elements involved may be related in a number of different ways. Or, in order to envisage possible alternative points of view with respect to a given set of data, the person may first have to conceive of, even at an intuitive level, the limits of a universe of possible relations that could form the basis for elaborating these alternative points of view. On the other hand, a minimum of analytic specification of the elements to be related may be necessary before the direction and the degree of their covariation can be specified.

The interpretation advanced here implies that the nuclear abilities of the capacities isolated should be considered the domain-oriented final crystallization of the factorially floating, but developmentally more basic, abilities. Clearly, more successful tasks than those used in this study have to be constructed if this presumably omnipresent capacity is to be directly scrutinized. It must be conceded that three (Planet 1, 2, and 3) out of the four items aimed at uncovering the earliest manifestations of the strategic capacity proved, in fact, to be tapping its very late manifestations. Therefore, tasks successful in this respect would only be those that were able to uncover the basic orientation and the processing characteristics of thought, which has been barely emancipated from the necessity to function under the guidance of perceptual, or personally familiar, reality. The authors are now intensively studying the thought of pre- and early adolescents, using a wide array of new tasks that have been de-

vised with these specifications in mind. However, the adequate understanding of structural relations is dependent upon at least equally adequate understanding of developmental relations.

The Development of Abilities

The results that concerned the development of the experimental and the relational sphere were in complete agreement with the third of the hypotheses stated. That is, the experimental capacity appeared to evolve from an almost mechanistic ability to produce combinations, and from a simple isolation-of-variables ability, to the ability to interconnect data so as to form valid working hypotheses that can be tested exhaustively until a coherent interpretative framework is constructed. The relational capacity begins its development as the ability to conceive of qualitative covariation, and is finally transformed into the ability to integrate ratio and proportionality quantitatively, after a person has acquired these two abilities in sequence. Therefore, the only deviation observed between the results and the relevant assumptions of STRATAMO was the synchronous, rather than sequential, acquisition of the ability to produce combinations and the first manifestation of the isolation-of-variables ability. However, this is not an important violation of STRATAMO's developmental assumptions. It simply indicates that, once thought becomes capable of cross-connecting elements so as to produce combinations, it is also capable of cross-connecting the properties (or levels) of two variables in order to isolate the one property from the other artificially.

The capacity to conceive of and reflect upon possibilities that was unexpectedly revealed by the results appeared to evolve from a very basic ability to derive conclusions from ready-made premises, by way of the ability to produce the premises themselves, to a postformal ability, which enables the subject to build up complementary or competing systems and to reflect on thought processes. The revelation of this capacity as a component of cognition that continues to develop after the appearance of the so-called tactical abilities enriches STRATAMO in one important respect: namely, it suggests that the strategic abilities of thought are not displaced by the tactical ones, once these have been acquired. On the contrary, they probably continue to pervade cognitive development and to exert significant influences on the various other, more domain-specific abilities. These conclusions are made all the more valid by the fact that, just as proposed under the fourth hypothesis, the developmental patterning of the abilities was the same across the various groups tested. Thus, the type of functional interdependence of the sequentially developing abilities must be specified.

Some of the developmental sequences described by Flavell (1972; also see Wohlwill, 1973) seem relevant in this respect. Specifically, a modification-type sequence seems appropriate to represent the emancipation of the various tactical abilities from the strategic capacity. In other words, some basic postconcrete *and* preformal abilities, whichever they may be, are transformed and refined by differentiation at some point in development, which should be placed around 14–15 years of age (see also, Kuhn, Ho, & Adams, 1979; Shayer & Adey, 1981; Webb, 1974).[3] This differentiation is aimed at matching special complex demands posed by reality (everyday and/or academic, as, e.g., the need to solve the complex algebraic or science problems posed at school). This match finally leads to the consolidation of the functionally different capacities that were uncovered by this psychological study.

However, once these capacities emerge as first formal-thought-step entities, they subsequently conform to both modification- and *inclusion*-type developmental relations. That is, an X ability (e.g., production of combinations; ratio or guided conception of the possible) is modified at some point in development to a Y ability (e.g., drawing hypotheses from combinations; proportionality or unguided conception of the possible, respectively), and both of them become included at some near point in development into an ability integrating both (e.g., production of new data yielding combinations in order to test the previously deduced hypotheses, and to arrive at a comprehensive theory; simultaneous application of ratio and proportionality in order to understand complex equilibrium systems, including both types of relations, or construc-

[3] Reference should be made here to some studies (Keating, 1975; Martorano & Zentall, 1980; Stone & Day, 1978) that obtained considerably higher proportions (up to 75%) of early adolescents succeeding in formal tasks. All of these studies examined bright subjects and used much fewer tasks than those used in this study. In addition, the tasks used by these studies were mainly addressed to the isolation-of-variables abilities, which were revealed by this study as being a relatively early achievement. Therefore, the results of this and the other studies are not actually in contradiction.

tion of complex complementary or competing ideal systems, respectively). It should be stressed here that these models are applicable to intrasphere developmental relations, that is, to those relations that were clearly validated by the high prediction values reported earlier.

A different model is needed to represent intersphere developmental relations, since, in their case, causal connections between abilities can only be indirectly inferred. For instance, the correlations observed between the various factors, though low or moderate, could be considered to indicate that the acquisition of the postformal capacity to conceive of and reflect upon various possibilities is positively affected by the previous acquisition of the third formal-step experimental and/or relational abilities. Indeed, this possibility seems very likely and is taken for granted by many authors (see Commons et al., 1982). Conversely, however, if the conceiving and reflecting capacity functions as a task-related, operations-scanning process, according to the conjectures advanced earlier, it may determine to a considerable extent how each of a series of tasks pertaining to the other capacities will be approached and processed. Hence, it may also determine the probability of success on the tasks, as well as the intertask consistency of performance. That is, it should be much more probable for a given pair of tasks to be processed by the same set of operations and solved at the same level, if they are judged similar, rather than different, under this process. This claim is valid and worthy of experimental scrutiny because of the coincidence observed between the subjective and the factorial structure of abilities.

However, the intersphere causal connections cannot be directly specified, as the prediction analysis has shown. In accordance with the third hypothesis, Flavell's (1972) *mediation* model seems to be the most appropriate to describe developmental relations of this type. The application of this model to the present context would imply that the acquisition of a preceding X ability, which belongs to one sphere, facilitates or even mediates the acquisition of a subsequent Y ability belonging to a different sphere. Nevertheless, in the case of mediation sequences, unlike the modification and inclusion sequences, X does not become an actual part or component of Y. Thus, once developed, Y functions independently of X, even though it has been mediated by it. This is the reason that, in the case of mediation sequences, the possible causal relations are not easily amenable to observation, especially if the data are cross-sectional. Longitudinal data collected by closely spaced testing intervals are badly needed if the intersphere developmental relations are to be sufficiently understood, particularly those concerned with the transformation of formal thought to postformal systematic and metasystematic thought (Commons et al., 1982), epistemic cognition (Kitchener, 1983), dialectical or relativistic thought (Kramer, 1983), or natural rationality (Kitchener & Kitchener, 1981).

The analysis of cognitive-developmental sequences attempted above deviates in an important respect from Piaget's assumptions regarding the processes that bring about cognitive change. Namely, it maintains that different stages (levels) or steps (substages, in Piagetian terms) of cognitive development differ not only in the quality of abilities characterizing each of them, but also in the quality of the processes that cause progress from the one to the other level or step. Consequently, the whole complex of abilities and processes is conceived of as a *cognitive-developmental loop.* That is, the abilities of a given step A, once they reach a certain degree of functional maturity, set in motion some *a* processes of change (e.g., of the modification type) that will cause the appearance of the abilities characterizing the next step B; these, in their turn, will cause the transformation of the *a* into *b* processes (e.g., of the inclusion type); once this transformation is effected, the B-step abilities will be reorganized and will take the form of the abilities characterizing the next step C, and so on. Evidently, this conception of cognitive development is in sharp contrast to the Piagetian conception, according to which the same and always present functional invariants (the gross concepts of assimilation and accommodation) cause the qualitative transformation of structures along the stage continuum described by Piaget (see Piaget, 1970). This dynamic interdependence between the processes and abilities of cognitive development has been experimentally documented and theoretically detailed by Demetriou (1983) in his studies of cognitive development that covered the age range 4–10.

Individual Differences

Age differences.—The differences among the three younger age groups are not especially interesting. That is, these differences came out as any developmental theory would have predicted. Specifically, the subjects of the three younger age groups showed, overall, an orderly progression from one age level to the next in all three capacities. An

exception to this orderly progression was the inhibition of development observed between the fifteenth and the eighteenth year of age in the conceiving and reflecting capacity. However, this fact need not be considered as destructive to developmental theory, given the massive representation of postformal abilities in this capacity. Postformal abilities seem to be very rare, even among university students, according to the results of this and other studies (Commons et al., 1982). Therefore, they could not be expected to be present in the secondary school adolescents.

Moreover, the diminution of correlations between factors observed in the case of the 18-year-olds might indicate that this age level is marked by an active reintegration of abilities that will lead from the mainly preformal abilities of the previous age level to the formal and postformal abilities of the next age level (see Wohlwill, 1973, for the possible meaning of the across-age-levels variation of the height of correlations). Evidently, it is not yet understood how, if at all, the reintegration of abilities that takes place during the transitional phases of development negatively affects their functioning, until they are elevated to the next more complex and effective level of functioning.

However, the lower performance of the 45-year-old subjects compared to that of the 21-year-olds, in all three capacities investigated, requires further discussion. The performance of these middle-aged subjects was at about the same level as that of the 18-year-olds of high SES in the experimental and the relational sphere, and even lower in the sphere of conception and reflection. Their responses to the units nested under this sphere were characterized by a marked and obstinate refusal to consider alternative, "unrealistic" possibilities. For instance, a categorical "no, no life could exist on another planet" was the response given by almost all of these subjects to the four items comprising the planet task. It was as though their thought had lost its playfulness and creativity.

It should be noted, however, that the method of this study does not enable one to decide whether this finding should be considered indicative of a developmental phenomenon or of an individual-difference phenomenon. Specifically, the complete lack of information about the cognitive-developmental history of these subjects hinders one from asserting whether or not they had lost previously acquired formal and postformal abilities. Comparable results could have been collected if a sample of middle-aged subjects had been examined who had the same level of education as the 21-year-olds in this study. In this case, no decline of performance should have been observed, if there had been no loss of previously acquired formal and postformal abilities. It is only on the basis of such information that developmental regression could be validly asserted. Otherwise one could justifiably assume that the lower performance of older-aged subjects is simply due to the fact that development never reached its end state in this particular group of subjects.

However, it should be noted here that the net finding of the middle-aged subjects' lower performance, especially in the capacity to conceive of and reflect upon various possibilities, should make researchers in this field particularly careful in asserting that relativistic or dialectical thinking (see Blasi & Hoeffel, 1974; Riegel, 1973) can be attained despite possible shortcomings in formal-thought functioning. After all, as Kitchener and Kitchener (1981, p. 179) have correctly pointed out, "it took humankind centuries before they even developed this relatively simple ability, and we should not depreciate or minimize it."

Sex differences.—The gender-differential results came out as predicted by hypothesis 5. Specifically, males performed better than females only in the relational sphere. This sex difference should be considered as simply quantitative, according to Buss's (1974) arguments, since no important structural difference was revealed by the relevant group-specific factor analyses. However, the difference is still interesting in two respects. First, it indirectly implies that the various thought capacities may be *differentially* influenced by internal and/or external factors, because they are not functionally equivalent. Such an interpretation, if pushed to its limits, leads to the following conclusion: there are not only strong and weak universals in cognitive development, as Dasen (1977) maintained, but also strong and weak differentials. That is, there are intergroup differences in some capacities that are less amenable to extinction than differences in other capacities. The distinction is based on the assumption that some capacities are more dependent on internal cognitive processes than others, which, to a larger extent, are based on externally available boosters. One could contrast here the mental rotation of covariation relations needed to grasp the direction of a given proportional relation to the verbal flexibility needed to formulate precise hypotheses, in the case of the experimental capacity (see De-

metriou & Efklides, 1981). This analysis does not seem unjustified, given that the sex differences observed persisted only in the relational sphere, even though the educational and professional experiences of the males and females tested were, in principle, identical across age and SES. Therefore, the controversy found in the literature over the problem of existence or nonexistence of sex differences in formal thought (Neimark, 1975) might be ascribed to the fact that those studies that found differences used relational tasks, where the two sexes are more prone to differ (e.g., Demetriou & Efklides, 1979), whereas the studies that did not find any consistent differences used only experimental tasks, where the sexes do not differ (Overton & Meehan, 1982; see also Shayer & Williams, 1983).

Second, one should examine the possibility that intergroup differences in capacities that fall within the weak-differential category could also be observed. However, differences of this kind could only be present insofar as a weak-differential capacity is dependent upon a strong-differential capacity. For example, females' performance could be lower than that of males, even in the experimental tasks, for which they functionally possess the directly relevant abilities. This difference may be explained and removed by reference to the fact that, according to the analysis advanced before, the effective application of experimental abilities is partially dependent on the application of some relational abilities. It should be mentioned here that the application of simple ANOVAs (not presented) on the factor scores showed that the performance of females was lower than that of males even in the experimental capacity. However, when the effect of the relational capacity on the experimental capacity was removed by means of covariance analysis, females' inferiority in the experimental capacity disappeared.

SES differences.—The effect of SES on the various capacities was as predicted by hypothesis 6. Namely, the high-SES subjects performed better than the subjects coming from rural areas in all three capacities; the performance of the low-SES came in between, in one case matching the performance of the high-SES subjects (experimental sphere), and in another case matching the performance of the rural subjects (relational sphere). These differences should be considered quantistructural, rather than simply quantitative (Buss, 1974). This assumption may be justified by the fact that the three groups differed both in structural cohesiveness, as indicated by the number of factors abstracted from their performance, as well as by the height of the correlations between factors, and in level of attainment, as indicated by the joint application of discrimination-level analysis and the analysis of the factor scores attained by the three groups. Therefore, one may conclude that, in relation to the fourth hypothesis, not only the attainment rates, but also the structural integration of presumably equivalent abilities, could be negatively affected by adverse environmental conditions.

The authors are not aware of any study that provides findings directly linking specific aspects of formal-thought structuring and developmental patterning to specific aspects of social or physical milieu. Therefore, the cause of the differences specified above is difficult to ascertain. Nevertheless, some theorizing and evidence concerning cognitive functioning and milieu connections in general might help future research on formal thought to focus on these crucial questions. Goodnow's ideas seem relevant to the problem of structuring. She said that "different cultures or subcultures may vary a great deal in the extent to which they emphasize working on complete answers, covering both present and possible events, as against the more pragmatic goals of locating only the information needed for the moment. They may also vary in the value and the practice given to finding loose ends, or coming up with questions" (1976, p. 174). This statement is in line with Bernstein's (1971) controversial and not widely accepted findings that show that the parent-child interaction of the lower social class is characterized by a restricted linguistic code, as contrasted to the elaborate code characterizing the interaction of the middle- and upper-social-class members. With regard to attainment rates, Bee, Van Egeren, Streissguth, Nyman, and Leckie (1975) have shown that the superiority of middle-class children's performance in cognitive tasks can be attributed to the fact that middle-class parents provide their children with more complex problem-solving strategies, compared to the strategies provided by the parents at the lower end of the SES hierarchy. Since linguistic proficiency enables the subject to interconnect and manipulate information at a high level of abstraction, it is related to formal thought (Markoulis, 1983). Therefore, insofar as formal-thought abilities require the invention and realization of complex problem-solving strategies, one may expect to observe the structuring and attainment gradation found in this study.

The above proposed notion of a cognitive-developmental loop helps one to understand the interdependence of structuring and level of attainment observed in the three SES groups. That is, an initial loose coordination (e.g., that observed in the rural subjects) that keeps separate even those abilities that belong to the same structure prevents the activation of the corresponding change processes that would lead to the attainment of the next level or step of development. In their turn, these processes, not having been activated, cannot bring about the amalgamation of the abilities that could eventually be integrated into a more general capacity, and so on. Therefore, development is trapped in a vicious circle that, if not broken by some sort of relevant benevolent influences, is destined to have a mostly undesirable result: that is, the continuous enlargement of the nonprivileged groups' disadvantage. This is probably the explanation for the fact that the rural subjects were functioning at the concrete level across all three capacities, even at the age of 18.

Finally, the notion of strong versus weak differentials that has been employed here to interpret the sex differences observed is also relevant to the interpretation of the SES groups' differences in the rates of attainment in the three capacities. Specifically, the relational capacity was considered to be a case of strong differentials, and it appeared to be the most insensitive to positive social/physical external influences. This seems to be why it was not affected by the demands posed by life in the complex urban environment. As a consequence, the low-SES subjects performed at the same low level as the rural subjects in this capacity. On the other hand, the experimental capacity proved to be the most sensitive to these demands, as the low-SES subjects performed at the same level attained by the high-SES subjects, despite the presence of obvious differences in their developmental histories. These demands also appeared to influence positively, though to a lesser extent, the capacity for conceiving of and reflecting upon various possibilities. As previously mentioned, the performance of the low-SES subjects was lower than the performance of the high-SES subjects in this capacity. It was higher, however, than the performance of the rural subjects. Therefore, the placement of the three capacities along the strong-to-weak differentials continuum should be as follows: relational → possibility conceiving-reflective → experimental capacity.

Overall, this study provided considerable support for most of STRATAMO's structural and developmental assumptions. Where a deviation was observed between assumptions and results, it extended and enriched the model, rather than invalidated it. Thus, the model appeared flexible enough to be integrated with recently developed conceptions regarding the structuring (Churchman's inquiring systems) and developmental sequencing (Flavell's developmental sequences) of complex thought abilities. In turn, the model provided further information about these conceptions and gave them a more precise form. That is, the model's basic assumptions and the results generated made it possible to conjecture that the organization and development of different thought capacities may be governed by different structural/developmental networks. These networks are the cornerstones of cognitive competence in general, that is, of a complex adaptive system that meets the different demands posed by the various domains of the environment and/or by the various phases of development, by constructing domain-specific structures and phase-specific processes. The model also appeared to be capable of being integrated with the current literature concerning cultural and individual differences. Therefore, further and finer-grained theoretical and empirical investigations within the context of this model could yield interesting answers to the problems raised in this discussion. These answers may finally lead to the rejection of STRATAMO, but, nevertheless, they will remain as part of our knowledge and understanding.

References

Bee, H. L., Van Egeren, L. F., Streissguth, A. P., Nyman, B. A., & Leckie, M. S. (1975). Social class differences in maternal teaching strategies and speech patterns. In U. Bronfenbrenner & M. A. Mahoney (Eds.), *Influences on human development* (pp. 297–308). Hinsdale, IL: Dryden.

Bernstein, B. (1971). Social class, language and socialization. In D. R. Dale, G. M. Esland, D. Mackinon, & D. F. Swift (Eds.), *School and society: A sociological reader* (pp. 139–147). London: Routledge & Kegan Paul.

Blasi, A., & Hoeffel, E. C. (1974). Adolescence and formal operations. *Human Development*, **17**, 344–363.

Broughton, J. M. (1981). Piaget's structural developmental psychology: 2. Logic and psychology. *Human Development*, **24**, 195–224.

Bruner, J. S. (1974). Going beyond the information given. In J. M. Anglin (Ed.), *Beyond the information given* (pp. 218–238). London: Allen & Unwin.

Burstein, B., Bank, L., & Jarvik, L. F. (1980). Sex differences in cognitive functioning: Evidence, determinants, implications. *Human Development*, **23**, 289–313.

Buss, A. R. (1974). Multivariate model of quantitative, structural and quantistructural ontogenetic change. *Developmental Psychology*, **10**, 190–203

Churchman, C. W. (1971). *The design of inquiring systems: Basic concepts of systems and organization.* New York: Basic.

Commons, M. L., Miller, P. M., & Kuhn, D. (1982). The relation between formal operational reasoning and academic course selection and performance among college freshmen and sophomores. *Journal of Applied Developmental Psychology*, **3**, 1–10.

Commons, M. L., Richards, F. A., & Kuhn, D. (1982). Systematic and metasystematic reasoning: A case for levels of reasoning beyond Piaget's stage of formal operations. *Child Development*, **53**, 1058–1069.

Dasen, P. (1977). Are cognitive processes universal? A contribution to cross-cultural Piagetian psychology. In N. Warren (Ed.), *Studies in cross-cultural psychology* (pp. 155–201). London: Academic Press.

Dasen, P., & Heron, A. (1980). Cross-cultural tests of Piaget's theory: Selected issues. In H. C. Triandis (Ed.), *Handbook of cross-cultural psychology* (pp. 113–159). New York: Allyn & Bacon.

Demetriou, A. (1983). Psychological development of the structures of concrete thought: Experimental studies on the thought of children aged from 4 to 10 years. *Scientific Annals of the School of Philosophy* (Supplement No. 39). Thessaloniki: Aristotelian University of Thessaloniki.

Demetriou, A., & Efklides, A. (1979). Formal operational thinking in young adults as a function of education and sex. *International Journal of Psychology*, **14**, 141–253.

Demetriou, A., & Efklides, A. (1981). The structure of formal operations: The ideal of the whole and the reality of the parts. In J. A. Meacham & N. R. Santilli (Eds.), *Social development in youth: Structure and content* (pp. 20–46). Basel: S. Karger.

Ennis, R. H. (1978). Conceptualization of children's logical competence: Piaget's propositional logic and an alternative proposal. In L. S. Siegel & C. J. Brainerd (Eds.), *Alternatives to Piaget: Critical essays on the theory* (pp. 201–260). New York: Academic Press.

Flavell, J. A. (1972). An analysis of cognitive-developmental sequences. *Genetic Psychology Monographs*, **86**, 279–350.

Froman, T., & Hubert, L. J. (1980). Application of prediction analysis to developmental priority. *Psychological Bulletin*, **87**, 136–146.

Goodnow, J. J. (1976). The nature of intelligent behavior: Questions raised by cross-cultural studies. In L. B. Resnic (Ed.), *The nature of intelligence* (pp. 169–188). Hillsdale, NJ: Erlbaum.

Goodnow, J. J., & Bethon, G. (1966). Piaget's tasks: The effect of schooling and intelligence. *Child Development*, **37**, 573–582.

Guilford, J. P. (1954). *Psychometric methods.* New York: McGraw-Hill.

Harman, H. H. (1967). *Modern factor analysis.* Chicago: University of Chicago Press.

Hollos, M., & Cowan, P. A. (1973). Social isolation and cognitive development: Logical operations and role taking abilities in three Norwegian social settings. *Child Development*, **44**, 630–641.

Humphreys, L. G. (1976). A factor model for research on intelligence and problem-solving. In L. R. Resnick (Ed.), *The nature of intelligence* (pp. 329–339). Hillsdale, NJ: Erlbaum.

Inhelder, G., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic.

Jamison, W. (1977). Developmental inter-relationships among concrete operational tasks: An investigation on Piaget's stage concept. *Journal of Experimental Child Psychology*, **24**, 235–253.

Joreskog, K. G., & Sorbom, D. (1978). *LISREL user's guide.* Chicago: National Educational Resources.

Keating, D. P. (1975). Precocious cognitive development at the level of formal operations. *Child Development*, **46**, 276–280.

Keating, D. P. (1980). Thinking processes in adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 211–246). New York: Wiley.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development*, **26**, 222–232.

Kitchener, K. S., & Kitchener, R. F. (1981). The development of natural rationality: Can formal operations account for it? In J. A. Meacham & N. R. Santilli (Eds.), *Social development in youth: Structure and content* (pp. 160–181). Basel: S. Karger.

Kofsky, E. (1966). A scalogram study of classificatory development. *Child Development*, **37**, 191–204.

Kramer, D. A. (1983). Post-formal operations? A need for further conceptualization. *Human Development*, **26**, 91–105.

Kuhn, D. (1979). *Intellectual development beyond childhood.* San Francisco: Jossey-Bass.

Kuhn, D., Ho, V., & Adams, C. (1979). Formal reasoning among pre- and late adolescents. *Child Development*, **50**, 1128–1135.

Markoulis, D. (1983). The impact of linguistic ability on cognitive development: A Neo-Piagetian

approach. *Scientific Annals of the School of Philosophy* (Supplement No. 35). Thessaloniki: Aristotelian University of Thessaloniki.

Martorano, S. H., & Zentall, T. R. (1980). Children's knowledge of the separation of variables concept. *Journal of Experimental Child Psychology*, **30**, 513–526.

Meacham, J. A., & Santilli, N. R. (Eds.). (1981). *Contributions to human development: Vol. 5. Social development in youth: Structure and content.* Basel: S. Karger.

Neimark, E. D. (1975). Intellectual development during adolescence. In J. Horowitz (Ed.), *Review of child development research* (Vol. **4**, pp. 541–594). Chicago: University of Chicago Press.

Neimark, E. D. (1979). Current status of formal operations research. *Human Development*, **22**, 60–67.

Overton, W. F., & Meehan, A. M. (1982). Individual differences in formal operational thought: Sex role and learned helplessness. *Child Development*, **53**, 1536–1543.

Peluffo, N. (1964). La nozione di conservazione del volume e le operazioni di combinazione come indici di sviluppo del pensiero operatorio in soggesti appartenenti ad ambienti fisici e socioculturali diversi. *Rivista de Psicologia Sociale*, **11**, 99–132.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (pp. 703–732). New York: Wiley.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development*, **15**, 1–12.

Richards, F. A., & Commons, M. L. (1984). Systematic, metasystematic and cross-paradigmatic reasoning: A case for stages of reasoning beyond formal operations. In M. L. Commons & C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 32–119). New York: Praeger.

Riegel, K. F. (1973). Dialetic operations: The final period of cognitive development. *Human Development*, **16**, 346–370.

Shayer, M. (1978). *A test of the validity of Piaget's construct of formal operational thinking.* Unpublished doctoral dissertation, University of London.

Shayer, M., & Adey, P. (1981). *Towards a science of science teaching: Cognitive development and curriculum demand.* London: Heineman Educational Books.

Shayer, M., Demetriou, A., & Prevez, A. (1983). *A four-cultures study of concrete operations: A fine-grained analysis of the stage concept.* Unpublished monograph.

Shayer, M., & Williams, J. (1983). *Sex differences on Piagetian formal operations tasks: Where they went and how to find them.* Manuscript submitted for publication.

Smedslund, J. (1964). Concrete reasoning: A study of intellectual development. *Monographs of the Society for Research in Child Development*, **29**(2, Serial No. 93).

Stone, A. C., & Day, M. C. (1978). Levels of availability of a formal operational strategy. *Child Development*, **49**, 1054–1065.

Thurstone, L. L. (1957). *Multiple-factor analysis.* Chicago: University of Chicago Press.

Tomlinson-Keasey, C., & Keasey, C. B. (1974). The mediating role of cognitive development in moral judgment. *Child Development*, **45**, 291–298.

Webb, R. A. (1974). Concrete and formal operations in very bright 6–11-year-olds. *Human Development*, **17**, 292–300.

Wohlwill, J. F. (1973). *The study of behavioral development.* New York: Academic Press.

Wood, P. K. (1983). Inquiring systems and problem structure: Implications for cognitive development. *Human Development*, **26**, 249–265.